LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

numero 175
marzo 1983
anno XVI
volume XXX

Coltivazione su grande scala di cellule di mammifero

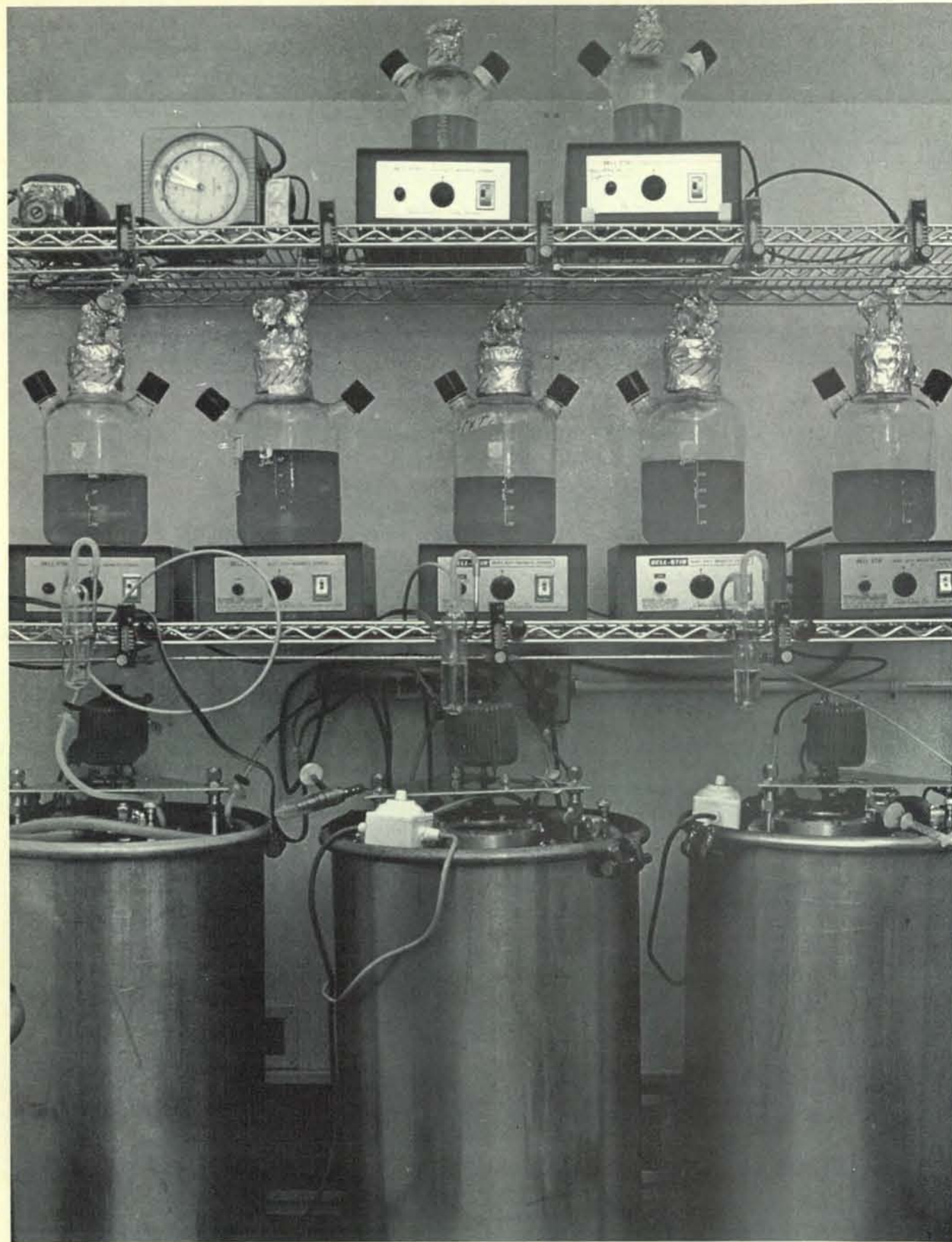
Sono in corso di sperimentazione nuovi reattori per la coltura di grandi quantità di quelle cellule fragili e complesse che sintetizzano proteine di notevole valore medico come l'interferon e gli anticorpi monoclonali

di Joseph Feder e William R. Tolbert

Vi sono molecole, di notevole valore dal punto di vista clinico, di studio e probabilmente commerciale, che possono essere prodotte più vantaggiosamente facendo crescere in coltura cellule umane o di altri mammiferi piuttosto che ricorrere a una produzione per via sintetica. Non è però semplice coltivare grossi quantitativi di cellule di mammifero in un mezzo artificiale. La tecnologia della microbiologia industriale, che già si trova a un buon livello di sviluppo, è adeguata alle esigenze di batteri, lieviti e muffe. Ogni singola cellula, racchiusa da una resistente parete cellulare, è una fabbrica metabolica indipendente che presenta richieste nutritive abbastanza semplici; nel caso dei batteri, sono spesso sufficienti glucosio e alcuni sali. I microrganismi crescono bene quando sono in grado di muoversi liberamente in un terreno di coltura liquido, in recipienti anche con una capacità di circa 275 000 litri; in queste condizioni sono in grado di resistere anche quando la proliferazione cellulare dà origine a una sospensione densa e questa venga sottoposta a un vigoroso mescolamento da parte di un agitatore meccanico.

Le cellule di mammifero sono diverse. Presentano dimensioni superiori, sono più fragili e più complesse rispetto a quelle della maggior parte dei microrganismi. Il delicato ectoplasto che circonda una cellula animale non è racchiuso in una parete cellulare resistente. I fabbisogni nutritivi della cellula di mammifero sono più rigorosi di quelli della maggior parte dei microrganismi e, infatti, non sono stati ancora completamente definiti. Più che essere un organismo libero, una cellula di mammifero è adatta a una vita specializzata come parte di un tessuto organizzato, il quale dipende da funzioni specializzate di molte altre cellule e da un sistema circolatorio che assicura a ogni cellula un ambiente perfettamente adatto e stabile. Una cellula di questo tipo è in grado di resistere quando viene separata dal suo tessuto e può crescere in un mezzo di coltura artificiale. La maggior parte delle cellule animali non crescerà affatto in sospensione, ma lo farà solo quando potrà aderire a una superficie. Nel corso degli anni sono state sviluppate delle tecniche per far crescere cellule di mammifero, su piccola scala, in laboratorio. Molto più difficile si è dimostrato il tentativo di farle crescere efficacemente su scala anche moderatamente superiore.

La necessità di disporre di metodi di coltura su grande scala è pressante. Consideriamo l'interferon; esso viene liberato da cellule animali e inibisce l'infezione virale. Pur essendo stato scoperto nel 1957, non è ancora stata stabilita con precisione la sua efficacia clinica principalmente perché è stato difficile far crescere in gran quantità le cellule di mammifero che possono essere indotte a sintetizzarlo. Per lo stesso motivo, gli anticorpi monoclonali specifici per una determinata proteina sono prodotti per mezzo di «ibridomi», formati dalla fusione (o ibridazione) di una cellula, che secerne anticorpi, con una cellula del mieloma maligno. Si possono far crescere, con una certa efficienza, ibridomi di topo come tumori, in animali da laboratorio. Gli anticorpi monoclonali umani promettono di essere potenti agenti terapeutici, a patto che si riesca a far crescere in coltura gli ibridomi umani che li producono.

L'urochinasi umana, un enzima che attivando il plasminogeno permette la formazione di plasmina, un enzima che scioglie i grumi di sangue, è ovviamente dal punto di vista clinico una promessa. Una buona fonte di urochinasi sono le cellule di rene umano, a loro volta difficili da far crescere su grande scala. I fattori angiogeni sono mediatori cellulari che promuovono la crescita dei vasi sanguigni. Sembra che i fattori angiogeni per uno specifico tumore, liberati da cellule tumorali, promuovano la vascolarizzazione e la crescita di tumori solidi. Enormi quantitativi di questi fattori sono indispensabili per ricerche che possono portare allo sviluppo di farmaci in grado di regolare l'angiogenesi (e probabilmente per controllare la crescita di un tumore); così come sono utili grandi quantitativi da somministrare direttamente per avviare la rivascolarizzazione. Le cellule umane portano antigeni di superficie: molecole che servono da marcatori per distinguere cellule «proprie» da cellule «non-proprie» e cellule tumorali da cellule normali. Gli antigeni tumorali sono l'obiettivo

Questi recipienti per la coltura di cellule di mammifero in grandi quantitativi si trovano nel laboratorio degli autori presso la Monsanto Company. I tre serbatoi cilindrici d'acciaio sono reattori da 100 litri che servono per far crescere cellule che non hanno bisogno di aderire a una superficie e possono crescere in sospensione in un terreno di coltura. I recipienti di vetro visibili sulle mensole poggiano ciascuno su un mescolatore magnetico che pone in rotazione un elemento immerso nel mezzo di coltura. Le cellule, una volta cresciute, vengono inoculate in un recipiente di dimensioni maggiori. Il colore dei terreni di coltura (rosso) è dovuto alla presenza di un indicatore acido-base.

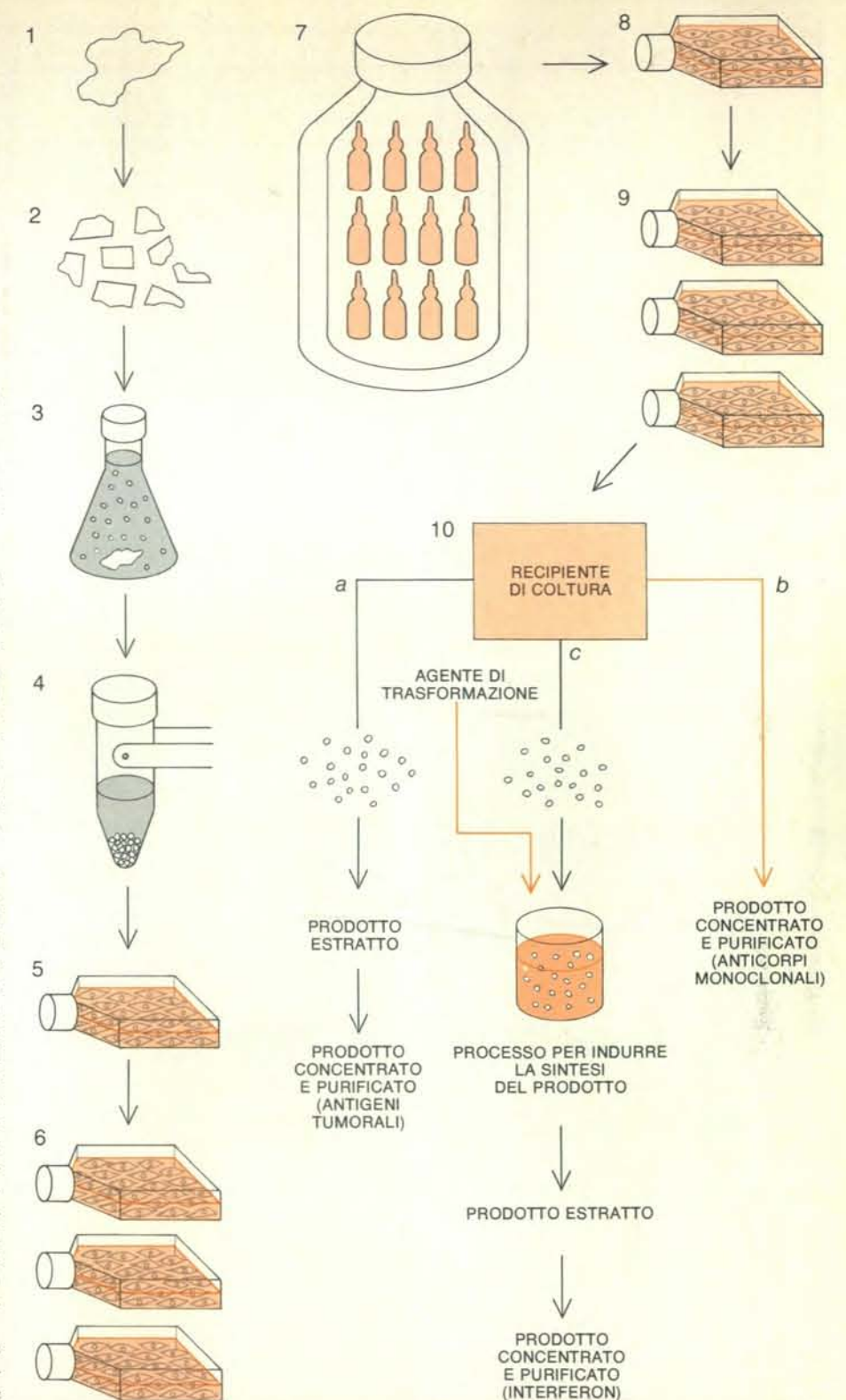
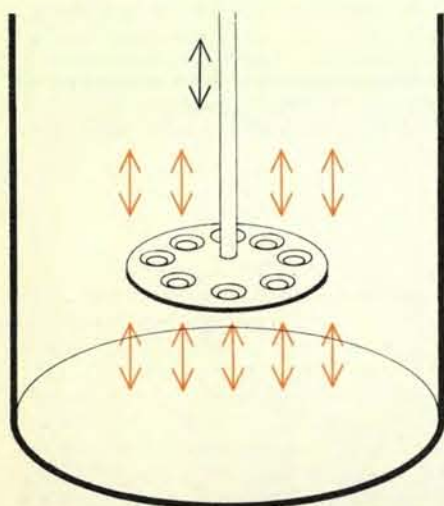
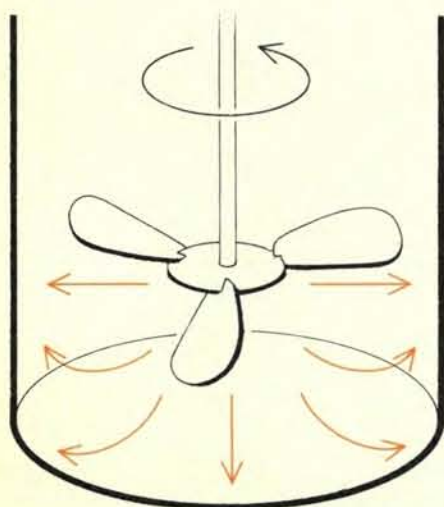
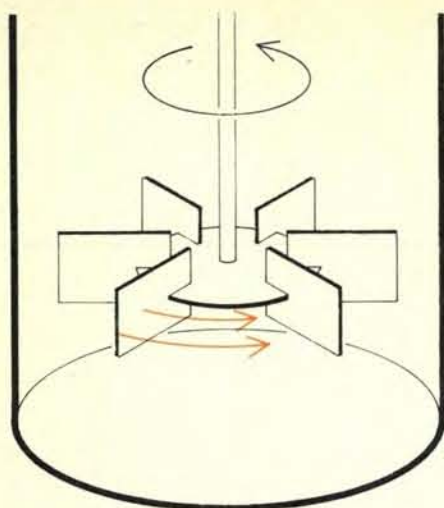
della maggior parte delle ricerche attuali e possono essere resi disponibili in quantità solo se si ricorre a una coltura di cellule su grande scala.

La richiesta di queste e molte altre molecole ha indirizzato gli sforzi recenti a sviluppare efficienti sistemi di coltura su grande scala per le cellule di mammifero. Gli sforzi non vengono scoraggiati dai notevoli successi ottenuti nella tecnologia del DNA ricombinante, che permette di introdurre il gene che codifica una data proteina nei batteri e quindi permette di indurre i batteri a sintetizzare la proteina. Molte molecole di importanza biologica non possono infatti essere sintetizzate in questo modo; nel caso poi di una sostanza, come il fattore angiogeno, che non è stata ancora neanche caratterizzata, la sola via per ottenere quantità sufficienti di materiale per procedere al suo isolamento è quella di far crescere le cellule che la producono.

Una coltura di cellule di mammifero parte da un tessuto. Questo viene dissociato meccanicamente o enzimaticamente, oppure tramite una combinazione dei due metodi, per ottenere una miscela di cellule singole e di piccoli ammassi cellulari. La miscela viene inoculata in un apposito terreno di coltura liquido che normalmente contiene sali, glucosio, alcuni amminoacidi (le subunità delle proteine) e siero sanguigno (che di solito rappresenta dal 5 al 20 per cento del terreno di coltura).

L'aggiunta di siero serve a fornire componenti che non sono stati ancora identificati, ma che si sono dimostrati necessari per la vita e la crescita delle cellule in un mezzo artificiale. Il siero è costoso e il suo costo determina, in larga misura, la fattibilità economica di un particolare sistema di coltura. Le cellule poste in coltura perdono l'elaborato sistema immunitario di difesa che è parte fondamentale dell'animale integro; per questo motivo si ricorre spesso all'aggiunta di antibiotici per prevenire la contaminazione della coltura. Bisogna, inoltre, controllare accuratamente il *p*H, la temperatura e la concentrazione di ossigeno e di anidride carbonica, così come va regolato il contenuto di sali per mantenere nei limiti la pressione osmotica del mezzo liquido, conservando pertanto l'integrità della fragile membrana cellulare.

Anche se la maggior parte delle cellule di mammifero deve aderire a una superficie solida, le cellule che hanno origine nel sangue o nel tessuto linfatico, insieme alla maggior parte delle cellule tumorali e ad altre cellule trasformate, possono adattarsi a crescere in sospensione. Queste cellule possono essere poste in coltura in apposite bottiglie di vetro, in cui il mescolamento è assicurato da un elemento che ruota grazie all'azione di un magnete esterno. Per applicazioni di laboratorio sono disponibili bottiglie di questo tipo con capacità che vanno da 25 millilitri fino a 15 litri.

Per le cellule che necessitano di una superficie di adesione (la maggior parte

Il procedimento generale mostrato in figura permette di produrre proteine di interesse medico avvalendosi di una coltivazione su grande scala di cellule di mammifero. Si taglia un tessuto *(1)* in piccoli pezzi *(2)*, che vengono trattati con un enzima proteolitico *(3)* il quale libera le singole cellule. Le cellule vengono raccolte dopo centrifugazione *(4)* e inoculate in un terreno nutritivo, dove proliferano, fino a coprire il fondo della beuta di coltura *(5)*. Le cellule confluenti vengono asportate dalla parete del recipiente per mezzo di un enzima e vengono reinoculate in numerose beute per estendere la coltura *(6)*. Alcune colonie di cellule vengono congelate in azoto liquido *(7)*. Alcune cellule vengono asportate in quantità sufficiente a inseminare una nuova coltura *(8)*, che viene estesa *(9)* per ottenere abbastanza cellule da inoculare in un recipiente di coltura di grande scala *(10)*. Le proteine possono essere recuperate in vari modi. Le cellule possono essere raccolte e ciò permette di rimuovere e purificare un prodotto cellulare come gli antigeni di superficie di cellule tumorali *(a)*. Se il prodotto desiderato è secreto dalle cellule, come nel caso di anticorpi monoclonali, si può asportarlo dal terreno di coltura che viene pompato dal reattore mentre le cellule vengono trattenute *(b)*. Talvolta si richiede un passaggio in più, in cui le cellule in coltura sono trattate per indurre la sintesi di un prodotto desiderato, come l'interferon *(c)*.

Per mantenere le cellule uniformemente sospese nel terreno e per distribuire regolarmente i gas disciolti è necessaria l'agitazione. L'agitatore a turbina *(in alto)* è comune nei fermentatori in cui crescono i microrganismi, ma la sua rapida rotazione determina consistenti forze di taglio orizzontali che possono danneggiare le fragili cellule di mammifero. Un agitatore a forma di elica marina *(al centro)* procura un rimescolamento sia orizzontale sia verticale a una velocità di rotazione inferiore. Il vibromescolatore *(in basso)* oscilla verticalmente, aumentando il mescolamento verticale, migliorando l'aerazione del terreno di coltura e riducendo il danno apportato alle cellule.

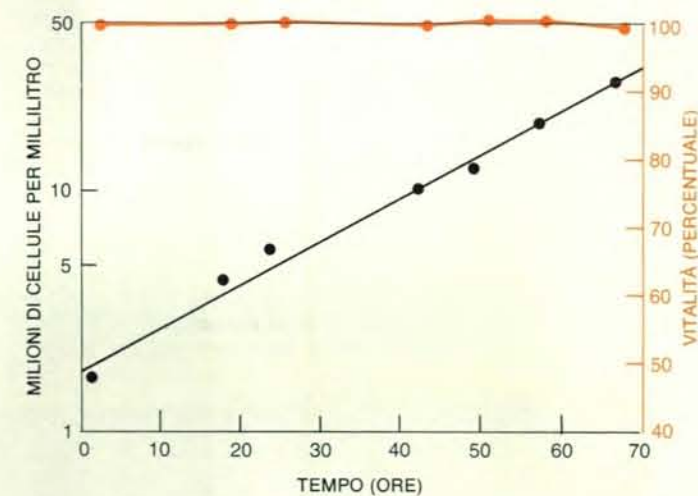
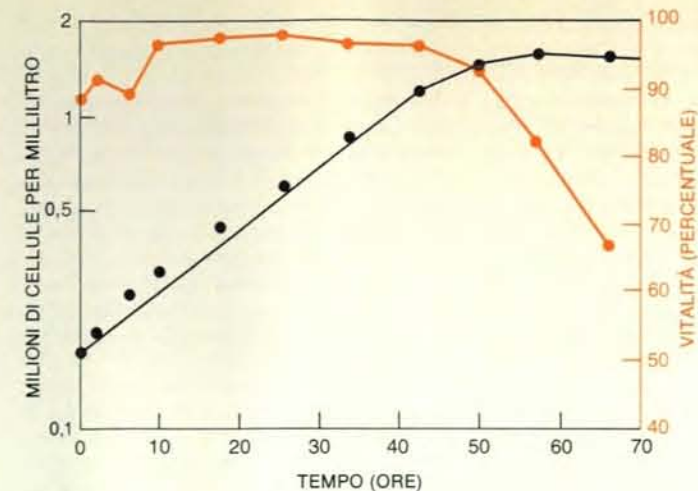
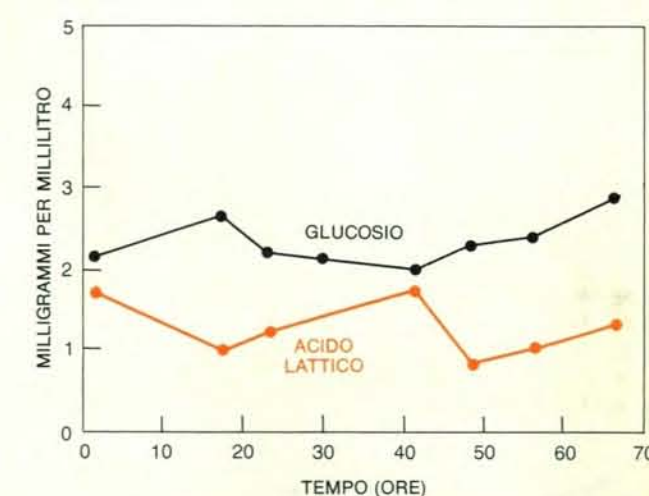
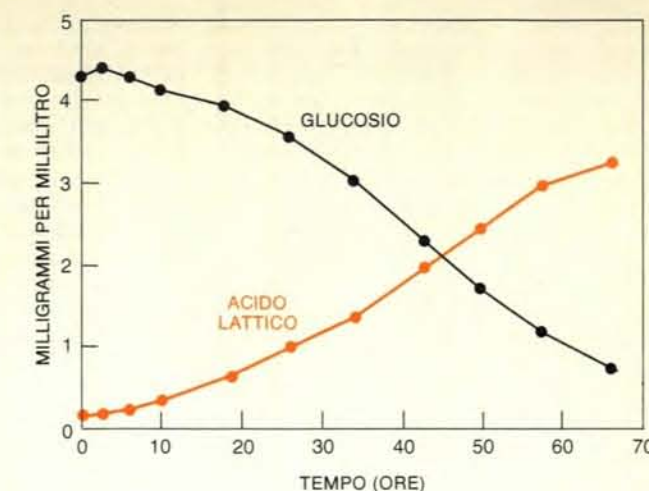
delle cellule normali) si possono impiegare diversi recipienti da laboratorio: piastre a micropozzo, piastre di Petri dal fondo piatto, e beute di varie dimensioni. Il recipiente più efficace è la bottiglia a cilindro, un recipiente cilindrico che viene orientato orizzontalmente ed è in rotazione continua. Le cellule si attaccano alla superficie interna del cilindro e il lento movimento di rotazione (circa un giro al minuto) le espone alternativamente al mezzo liquido e all'aria. Le bottiglie a cilindro più grandi mettono a disposizione una superficie di circa 1600 centimetri quadrati.

Per le cellule che crescono in sospensione sono stati approntati reattori di grande scala, che si basano su modifiche apportate ai fermentatori in cui vengono fatti crescere i microrganismi unicellulari. Il processo, tuttavia, presenta tuttora dei problemi. La maggior parte dei reattori derivati dai fermentatori è complessa e costosa. Per mantenere l'ambiente sterile essi fanno affidamento su un complicato dispositivo di sterilizzazione del vapore e sull'aggiunta di antibiotici al terreno di coltura. La crescita su grande scala di cellule che necessitano di una superficie di adesione presenta problemi ancora più difficili, che fino a poco tempo fa venivano superati moltiplicando semplicemente il numero di recipienti convenzionali, soprattutto bottiglie a cilindro.

Alcuni anni or sono il nostro laboratorio nel dipartimento di ricerca della Monsanto Company si è impegnato a ideare nuovi sistemi di coltura cellulare su grande scala in grado di ovviare alla necessità di una fonte sicura di centinaia di litri di sospensioni cellulari ogni settimana. All'inizio abbiamo deciso che i sistemi devono mantenere un'assoluta sterilità senza ricorrere ad antibiotici, che possono mascherare il lento insorgere di una modesta contaminazione per un certo periodo e possono altresì influenzare il metabolismo delle cellule in coltura in modo non sempre prevedibile. Abbiamo anche optato per un metodo modulare: i singoli recipienti di coltura dovevano essere di dimensioni ridotte in modo da ottenere una produzione su grande scala mettendo semplicemente più moduli in linea. Da ultimo abbiamo stabilito di studiare a fondo i progetti innovativi che ci consentivano di far fronte alle specifiche condizioni richieste dalla crescita delle grandi e fragili cellule di mammifero.

Tutti questi criteri per la crescita di cellule in sospensione furono soddisfatti da un reattore da 100 litri, semplice e poco costoso. Era costituito da un tamburo di acciaio inossidabile da 30 galloni (135 litri), munito di una testa asportabile che permetteva di eseguire tutte le aggiunte e tutti i prelievi. Il reattore è sufficientemente piccolo da essere sterilizzato in autoclave; un portello a doppio setto, situato nella testa asportabile, permette di prelevare i campioni tutte le volte che è necessario senza contaminare la coltura. L'agitazione richiesta viene fornita da un vibromescolatore (un dispositivo illustrato per la prima volta nel 1965 da George E. Moore del Roswell Park Memorial Institute). Un disco con aperture coniche è attaccato a un albero verticale che oscilla in su e in giù ad alta velocità. La vibrazione del disco e il passaggio del terreno di coltura attraverso i fori si combinano e l'esito delle due azioni combinate è una vigorosa agitazione del mezzo più in direzione verticale che non orizzontale. Un ulteriore vantaggio del vibromescolatore è dato dal fatto che il suo movimento viene trasmesso per mezzo di un diaframma flessibile che fornisce una tenuta assoluta e statica alla sommità del recipiente, attraverso la quale alcuni tubi fanno gorgogliare ossigeno nel terreno e muovono l'aria ricca di anidride carbonica sopra la superficie del mezzo di coltura.

Abbiamo costruito numerosi reattori di questo tipo da 100 litri a un costo di circa un ventesimo di quello di fermentatori convenzionali. Essi sono adatti a una gran varietà di cellule che crescono in sospensione e nel corso degli ultimi anni sono stati fatti crescere e quindi raccolti centinaia di migliaia di litri di sospensioni di cellule umane e di cellule di roditori. Ciononostante, abbiamo deciso di non costruire versioni più grandi del reattore da 100 litri per ottenere livelli produttivi ancora superiori. Abbiamo invece scelto di procedere a sviluppare un sistema che fosse in grado di imitare maggiormente le condizioni di crescita, in natura, delle cellule di mammifero.

Un sistema convenzionale di coltura cellulare, come il nostro recipiente da 100 litri, espone le cellule che stanno crescendo a condizioni ambientali in continua modificazione. Le cellule vengono inoculate in un mezzo che contiene una quantità appropriata di idonee sostanze nutritive ed è privo di prodotti metabolici di scarto; al crescere delle cellule, la concentrazione di sostanze nutritive diminuisce mentre si accumulano i prodotti di scarto. Questo fatto impone un limite alla densità della coltura. Negli animali viventi, d'altro canto, questa regolazione è assicurata dall'omeostasi. Un sistema circolatorio efficiente libera ossigeno e sostanze nutritive e rimuove anidride carbonica e prodotti di scarto, fornendo a ogni cellula un ambiente stabile, anche se la densità cellulare è da 500 a 1000 volte superiore in un tessuto vivente rispetto a quella di un tipico mezzo di coltura. Sarebbe auspicabile riuscire a fornire alle cellule in coltura un ambiente quasi costante; una possibilità è quella di immettere continuamente terreno nuovo nella coltura e, altrettanto costantemente, di rimuovere il terreno esaurito. L'abilità consiste nel rinnovare il mezzo di coltura senza rimuovere le cellule. Nel 1969 Philip Himmelfarb e Philip S. Thayer della Arthur D. Little, Inc, hanno sviluppato un sistema di perfusione, su scala di laboratorio, per colture in sospensione di cellule animali e hanno riferito di aver ottenuto densità cellulari molto superiori di quanto non sia possibile con sistemi in sospensione di tipo convenzionale.

L'efficienza del sistema a perfusione è stata dimostrata dal raffronto tra crescita cellulare *(a sinistra)* ed esaurimento del terreno di coltura *(a destra)* in una bottiglia convenzionale a mescolamento magnetico *(in alto)* e in un reattore a perfusione da quattro litri *(in basso)*. Cellule tumorali di ratto, che crescono bene in sospensione, sono proliferate circa allo stesso tasso in entrambi i sistemi, ma hanno raggiunto una densità cellulare molto superiore nel sistema a perfusione *(curve in nero a sinistra)*, conservando la loro vitalità anche a elevata densità *(curve in colore a sinistra)*. Le concentrazioni di glucosio e acido lattico riflettono l'esaurimento di sostanze nutritive e la comparsa di prodotti di scarto. La concentrazione di glucosio diminuisce costantemente e quella di acido lattico aumenta regolarmente nella bottiglia a mescolamento magnetico convenzionale; nel reattore a perfusione si conservano concentrazioni costanti di sostanze nutritive e prodotti di scarto.

Abbiamo realizzato un sistema di perfusione su media scala con recipienti da 4 e da 44 litri in cui le cellule vengono fatte crescere in quantità paragonabili a quelle ottenute normalmente in recipienti con capacità di 100 e 1000 litri. La perfusione è resa possibile dall'introduzione nel sistema di un recipiente satellite contenente un filtro cilindrico di porcellana con pori di un diametro inferiore a due micrometri (millesimi di millimetro). Il filtro trattiene le cellule che vengono ripompate nel recipiente di coltura; parte del terreno di coltura viene riciclata e parte viene asportata quando si alimenta il sistema con terreno nuovo. Per minimizzare l'intasamento dei fori del filtro da parte di cellule e frammenti cellulari, si esercitano alla sommità del filtro forze di taglio, facendo ruotare il filtro stesso oppure circondandolo con un agitatore rotante. Sensori speciali, portelli di campionamento e sistemi di tubi di immissione e prelievo permettono di mantenere il pH e la concentrazione di ossigeno, anidride carbonica e sostanze nutritive al livello più opportuno, anche quando le cellule proliferando hanno ormai raggiunto una densità molto elevata.

Il conseguimento di queste densità, grazie alla perfusione, ha reso necessaria la progettazione di un nuovo meccanismo di agitazione più blando. Quattro lamine flessibili ottenute da un monofilamento di plastica ruotano lentamente (da 8 a 20 giri al minuto) nel recipiente di coltura, gonfiandosi come vele quando il terreno di coltura liquido viene mosso dal loro bordo d'uscita. Distribuendo l'energia richiesta per l'agitazione su un'ampia superficie, il meccanismo a vela flessibile riduce le forze di taglio che agiscono sulle singole cellule.

Per valutare l'efficienza del sistema di perfusione abbiamo posto a confronto la crescita di cellule tumorali di ratto in questo e in un sistema convenzionale e abbiamo tenuto sotto controllo la concentrazione di glucosio e di acido lattico (un importante prodotto metabolico) nel terreno di coltura che ne usciva. Il tasso di crescita delle cellule era più o meno lo stesso nei due sistemi, ma la densità cellulare era molto superiore nel reattore a perfusione: circa 25 milioni di cellule per millilitro rispetto a un milione di cellule per millilitro nel reattore convenzionale. Alla massima densità tutte le cellule del sistema a perfusione erano vive, mentre circa il 30 per cento delle cellule nel sistema convenzionale erano morte. Cosa probabilmente ancora più significativa dal punto di vista economico, la resa di cellule per litro di terreno di coltura consumato era di più di due volte e mezzo superiore nel sistema a perfusione (anche se in un sistema di questo tipo il terreno viene continuamente rinnovato). Inoltre, abbiamo trovato che per il reattore a perfusione è possibile un modo di funzionamento chemostatico, in cui cioè insieme al terreno esaurito vengono prelevate in continuo anche alcune cellule, il che consente di mantenere una densità cellulare costante a un livello ottimale.

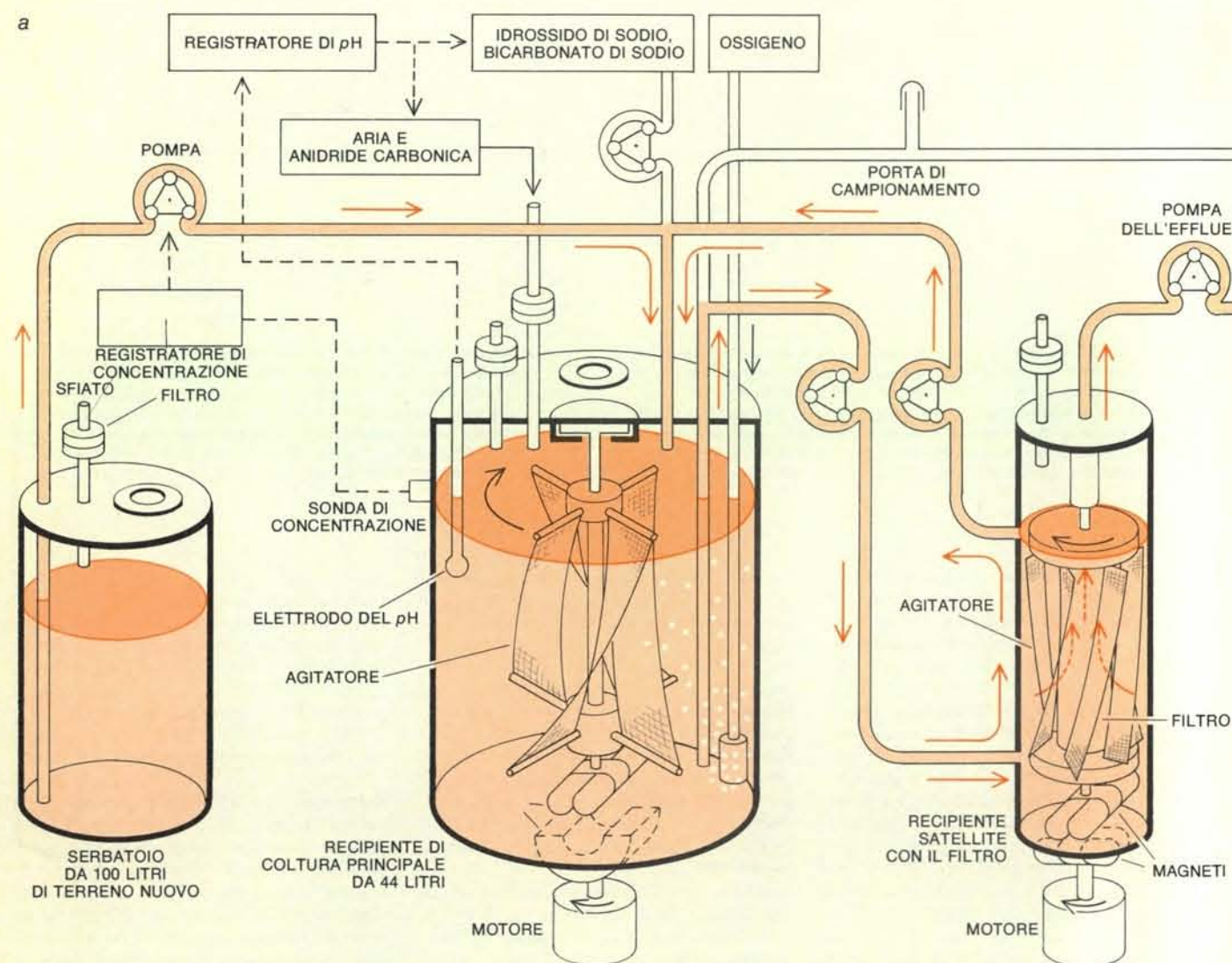
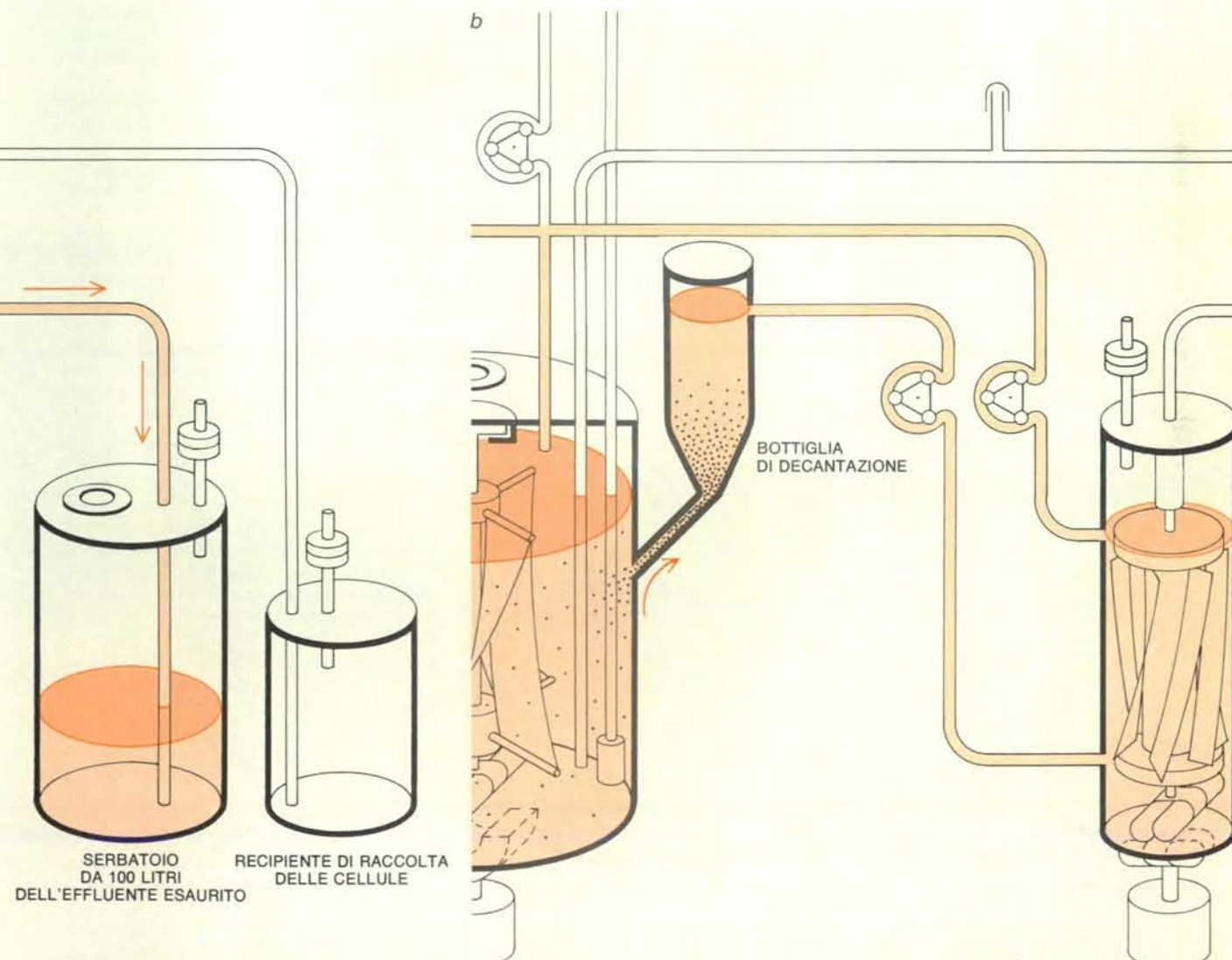
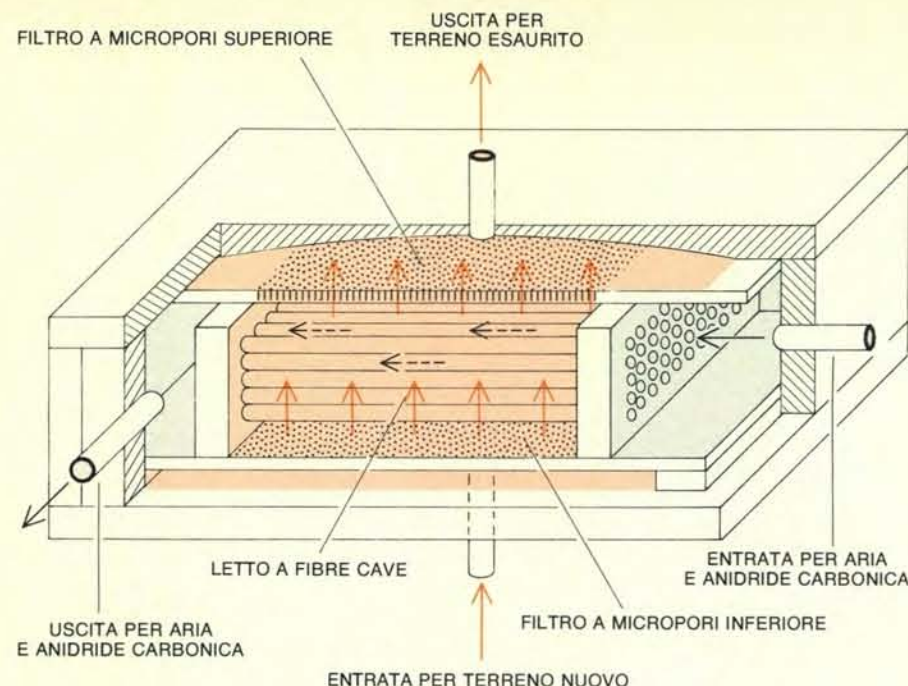
La grande maggioranza di cellule di mammifero normali, come abbiamo accennato in precedenza, cresce solo quando può aderire alla superficie del recipiente di coltura. Le bottiglie a cilindro e le beute non sono in grado di fornire un rapporto superficie/volume sufficientemente elevato per la coltura industriale su grande scala. Il rapporto superficie/volume può essere aumentato sfruttando vari stratagemmi. Per esempio, si possono far crescere le cellule su polimeri spugnosi, su schiere di sottili tubicini o fibre cave, su pile di sottili piastre o su microscopiche perline chiamate microtrasportatori. Proprio per applicazioni su grande scala abbiamo approfondito lo studio di due fra queste possibilità: le fibre cave e i microtrasportatori.

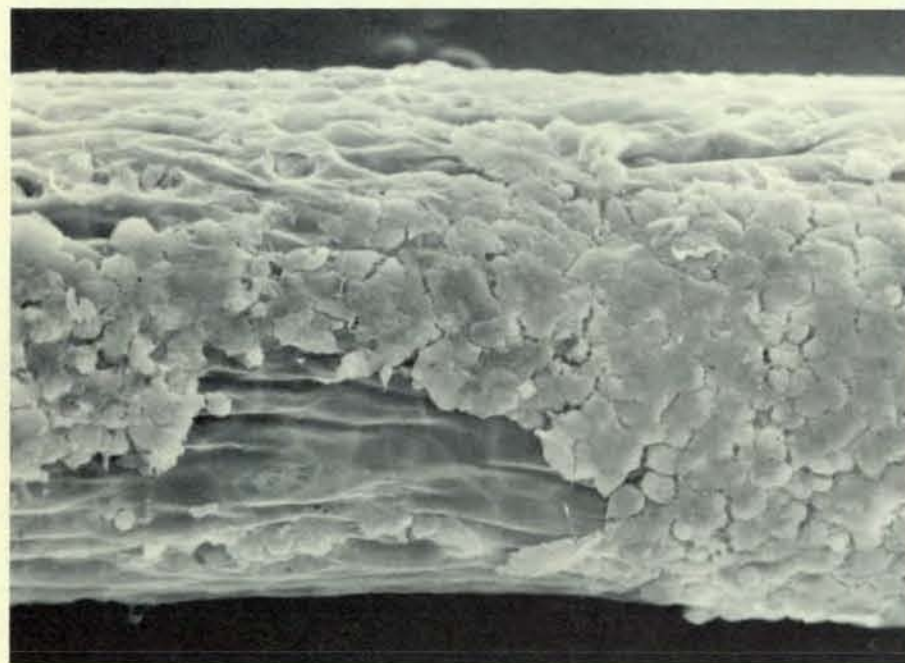
Studi eseguiti nel nostro laboratorio e da Richard A. Knazek e collaboratori al National Cancer Institute hanno dimostrato che cellule di mammifero crescono in modo soddisfacente su molte fibre cave con un diametro esterno compreso tra un terzo e tre quarti di millimetro. Le fibre sono porose, di modo che l'aria che le attraversa si diffonde alle cellule sulla superficie esterna, che è immersa nel terreno di coltura. Dopo aver studiato numerose configurazioni, alla fine abbiamo deciso per una disposizione a «letto piatto» delle fibre. Il nucleo del sistema è un letto di fibre estremamente sottile (da tre a sei strati) al quale si attaccano le cellule e dove proliferano.

Il terreno di coltura nuovo entra nel contenitore in acciaio inossidabile al di sotto del letto di fibre, diffonde attraverso il letto e viene prelevato alla sommità del contenitore. Affinché il flusso sia uniforme e perpendicolare al letto di fibre, il terreno di coltura viene pompato fino al letto attraverso un filtro a micropori: una lamina di acciaio inossidabile con pori di due micrometri di diametro. Un altro filtro, con pori da 20 micrometri, viene posto sopra al letto per limitare il ritorno di flusso. Il vantaggio della configurazione a letto sottile è che il percorso di perfusione è molto breve (rispetto, per esempio, al percorso in un sistema a «cartuccia» cilindrica, dove il flusso del terreno di coltura è parallelo all'asse delle fibre.) Di conseguenza è difficile che si formi un qualsivoglia gradiente da un elevato a un basso livello di sostanze nutritive e da un basso a un elevato livello di prodotti di scarto; praticamente tutte le cellule sono in presenza di terreno di coltura nuovo. Le estremità delle fibre sono legate assieme con un materiale elastico e non tossico, che viene tagliato in modo tale che le anime cave di ogni fibra si aprano nelle camere dove viene fatta circolare aria e anidride carbonica.

Abbiamo costruito due reattori prototipo, con una superficie di fibre rispettivamente di 930 e 9300 centimetri quadrati. (Il prototipo più grande misura solamente 40 centimetri quadrati e ha un'altezza di 4,5 centimetri.) Essi sono in grado di conservare colture ad alta densità per lunghi periodi di tempo. Per esempio, il prototipo più piccolo ha permesso la crescita di cellule di polmone umano con una densità di circa un milione di cellule per centimetro quadrato, cioè circa dieci volte la densità che può essere ottenuta in bottiglie a cilindro. In questi reattori la perfusione del mezzo è particolarmente efficace; può essere regolata per mantenere costante il livello di sostanze nutritive e di prodotti di scarto nel corso della proliferazione cellulare. Questo controllo può essere conseguito saggiando la concentrazione di acido lattico nel terreno di coltura esaurito, ottenendo così una stima del numero di cellule presenti sulle fibre e regolando di conseguenza il tasso di flusso del terreno nuovo.

Il sistema permette non solo la crescita di grandi quantitativi di cellule; in realtà costruisce un tessuto artificiale quando si ha la proliferazione di miliardi di cellule che ricoprono le fibre e arrivano persino a introdursi in esse. Grandi quantitativi di cellule vengono pertanto mantenuti in condizioni minimali di tensione per un lungo periodo di tempo (nei nostri esperimenti da 21 a 59 giorni), durante il quale vengono prodotte in continuo e raccolte dal terreno di coltura esaurito importanti molecole come l'urochinasi e il fattore angiogeno. Un attento controllo della contaminazione dovrebbe permettere di far funzionare uno di questi sistemi quasi indefinitamente. Una volta che è stata ottenuta la crescita massima è spesso conveniente sostituire il terreno (decisamente più costoso) che facilita la proliferazione cellulare con un semplice terreno di mantenimento, contenente poco o niente siero e badare solamente a mantenere in attività le cellule che producono una data molecola.

Il metodo che utilizza i microtrasportatori è stato illustrato per la prima volta nel 1967 da Anton L. van Wezel dell'Istituto nazionale olandese per la sanità pubblica e ne sono state studiate e sfruttate molte varianti per la produzione industriale su media scala. I microtrasportatori sono di destrano, un polimero naturale del glucosio, o di uno dei tanti polimeri di

Il reattore a perfusione, progettato dagli autori, immette nel sistema del terreno nuovo e asporta quello contenente le sostanze nutritive esaurite e i prodotti di scarto che si sono accumulati, conservando al tempo stesso le cellule all'interno del sistema. Il diagramma mostra la versione del reattore messa a punto per cellule che crescono in sospensione *(a)* e una versione modificata *(b)* per cellule che crescono su perline microscopiche chiamate microtrasportatori. Nel recipiente di coltura principale quattro fogli flessibili di nylon a monofilamento, che ruotano lentamente, forniscono una blanda agitazione; il terreno di coltura viene continuamente pompato dal recipiente principale al recipiente satellite contenente il filtro, dove parte del terreno viene pompata ancora, attraverso il filtro di porcellana, al recipiente dell'effluente. Il filtro trattiene le cellule, che vengono pompate, insieme alla parte rimanente del terreno, nel recipiente principale per un'ulteriore proliferazione. Una piccola versione dell'agitatore a vela flessibile ruota nel recipiente satellite per rendere minima l'otturazione dei pori del filtro. Le sostanze che vengono secrete dalle cellule possono essere raccolte dal terreno nel recipiente dell'effluente; le cellule possono essere pompate fuori dal reattore nel recipiente di raccolta delle cellule. La maggior parte delle normali cellule di mammifero cresce solo quando può aderire a una superficie, mentre non è in grado di crescere in sospensione in un terreno di coltura liquido; i microtrasportatori mettono a disposizione la superficie necessaria. Per poter adattare il reattore a perfusione alla crescita di cellule su microtrasportatori è stata posta tra il recipiente principale e quello satellite una bottiglia di decantazione. Le perline incrostate di cellule si dispongono all'esterno della sospensione, formando in tal modo un impasto liquido che scivola all'indietro nel recipiente principale, dove le cellule continuano a proliferare.

Il reattore a perfusione a fibre cave fa crescere le cellule su un letto piatto di fibre porose di plastica del diametro di 0,3 millimetri. Le estremità delle fibre sono aperte per consentire l'afflusso di aria e di anidride carbonica lungo le fibre *(frecce in nero)*; la parte esterna delle fibre è immersa in un terreno di coltura il cui flusso *(frecce in colore)* è mantenuto uniforme da due filtri a micropori. Le cellule inoculate nel reattore si attaccano alla superficie delle fibre e proliferano, formando col tempo quello che è in effetti un tessuto artificiale che può essere conservato per parecchi mesi.

sintesi, e presentano un diametro variabile da 50 a parecchie centinaia di micrometri. Essi sono sospesi in un terreno nutritivo; le cellule che necessitano di un rapporto di adesione, inoculate nel terreno di coltura, si legano alle perline e si moltiplicano. Il vantaggio di un sistema a microtrasportatori è dovuto al fatto che, nonostante le cellule crescano su una superficie, le perline che forniscono questo supporto sono sospese nel mezzo e di conseguenza le cellule possono essere fatte crescere in recipienti a sospensione convenzionali.

Questa microfotografia elettronica a scansione mostra cellule cancerogene umane (cellule HeLa) che crescono sulla superficie di una fibra cava. Alla fine, oltre a ricoprire l'intera superficie della fibra, le cellule penetrano nella sua struttura a lobi. L'ingrandimento è di circa 250 diametri.

Questi sistemi convenzionali non sono, tuttavia, molto efficienti per le colture che si avvalgono dei microtrasportatori. Oltre alle solite preoccupazioni, come quelle di controllare la concentrazione delle sostanze nutritive e dei prodotti di scarto, si aggiunge la necessità di considerare la quantità e la distribuzione dei microtrasportatori nel sistema. Le collisioni tra perline possono danneggiare le cellule. Poiché queste collisioni aumentano quando l'agitazione è elevata, si presenta la necessità di un meccanismo di mescolamento più delicato per mantenere le perline in sospensione senza distruggere troppe cellule. Per la crescita di grandi quantitativi di cellule è necessaria una superficie estesa, il che significa una elevata densità di perline; ma una densità troppo elevata comporta più collisioni. Il rapporto relativamente elevato superficie/volume ottenuto con i microtrasportatori dà origine a un ulteriore problema: man mano che le cellule proliferano può verificarsi un impoverimento molto rapido di sostanze nutritive e un parallelo arricchimento di prodotti di scarto tossici.

Da ultimo, vi è il problema di passare da un dato recipiente di coltura a un altro più capace. Le cellule incrostate devono essere rimosse dalle perline, dopo di che vengono sospese di nuovo nel recipiente più grande con una nuova aggiunta di perline. Nei sistemi a microtrasportatore convenzionali è stato difficile rimuovere le cellule dalle perline senza danneggiarne una grossa porzione.

Abbiamo adattato il nostro nuovo sistema a perfusione per utilizzare i microtrasportatori e abbiamo trovato che esso rendeva possibile un significativo aumento nella densità delle cellule. Gli agitatori a vela flessibile forniscono il delicato mescolamento necessario; facendoli girare da 8 a 20 giri al minuto, essi mantengono una elevata concentrazione di perline in sospensione nel mezzo (fino a 12 grammi di perline per litro di terreno di coltura) senza danneggiare le cellule. La principale modifica è stata l'aggiunta di una bottiglia di decantazione tra il recipiente del reattore principale e il recipiente satellite che alloggia il filtro. La pompa di riciclaggio muove il mezzo, con le sue perline in sospensione incrostate di cellule, verso il recipiente satellite. Una volta nella bottiglia di decantazione, rimosse dall'area dove il mezzo è sottoposto a mescolamento, le perline tendono a sedimentare e ad aggregarsi in un denso impasto che gradualmente scivola verso il recipiente principale; il mezzo, essenzialmente privo di perline, si muove verso il recipiente satellite che contiene il filtro e viene o riciclato o rimosso dal sistema.

Nel processo di decantazione le perline e le cellule a esse attaccate entrano in stretto contatto e continuano a essere perfuse con nuovo terreno di coltura proveniente dal recipiente principale. Le cellule si spostano da un microtrasportatore all'altro, formando grossi ammassi di perline e di cellule, con il risultato che si possono trovare fino a quattro volte più cellule di quante ci si aspetterebbe su una data superficie di perline. Da questa condizione di forte aggregazione deriva anche un altro vantaggio, che non avevamo previsto: la separazione delle cellule dai loro microtrasportatori è più facile e si risolve in questo modo il problema del trasferimento a un altro recipiente più grande. Nel corso dell'aggregazione si riduce leggermente l'adesione delle singole cellule alle perline controbilanciata però da adesioni multiple tra cellula e cellula. Quando l'ammasso di perline e di cellule viene aspirato dal reattore, è sufficiente un breve trattamento con l'enzima tripsina per liberare le cellule dai microtrasportatori in buone condizioni.

Il reattore a perfusione e microtrasportatori si è dimostrato un sistema efficace per la coltura su grande scala di cellule che necessitano di una superficie di adesione. I microtrasportatori hanno sempre avuto la capacità di mettere a disposizione una grossa superficie che può essere globalmente sfruttata (in realtà supersfruttata, dato l'instaurarsi di un effetto ponte tra le cellule) in quanto la perfusione conserva un ambiente ottimale per la crescita delle cellule, anche quando la densità cellulare raggiunge valori estremamente elevati (e in questo particolare sistema le densità cellulari sono da 5 a 10 volte superiori che in passato).

La combinazione della crescita cellulare ad alta densità e dell'aumento graduale dei recipienti di coltura porta a risultati impressionanti. In un ciclo di funzionamento sperimentale di un piccolo reattore (quattro litri) a perfusione e microtrasportatori abbiamo fatto crescere 40 miliardi di fibroblasti umani, equivalenti al quantitativo ottenuto normalmente da 1300 bottiglie a cilindro. Queste cellule sono state separate dai loro trasportatori e inoculate a loro volta in un reattore da 44 litri insieme a 400 grammi di nuovi microtrasportatori, che mettevano a disposizione una superficie di 188 metri quadrati. Nel ciclo di funzionamento in questione sono cresciuti 340 miliardi di cellule, una quantità che avrebbe richiesto 11 000 bottiglie a cilindro. I fibroblasti raccolti servivano non solo per la produzione di interferon, ma anche di urochinasi e di un fattore angiogeno. La misura migliore del rapporto costi/rendimento del sistema a perfusione e microtrasportatori è la resa di cellule per litro di terreno di coltura esaurito. In tutti i nostri esperimenti la resa è stata di circa quattro volte superiore a quella dei sistemi che utilizzano bottiglie a cilindro.

Per soddisfare le necessità dei laboratori di ricerca e della produzione su scala commerciale sono necessari efficienti sistemi di grande scala per la crescita di un'ampia gamma di cellule di mammifero, in assenza di antibiotici e a costi ragionevoli. Appare evidente come le nuove tecnologie possano ampliare la conoscenza dei processi biologici e fornire al tempo stesso quantità significative di prodotti cellulari, che promettono di diventare importanti agenti terapeutici.

Applicazioni biomediche del laser

Si basano sull'interazione tra la radiazione del laser e il campione biologico e permettono di realizzare effetti che vanno dalla microchirurgia cellulare alla chirurgia macroscopica dei tessuti

di Carlo A. Sacchi e Riccardo Pratesi

Il laser, in virtù delle sue caratteristiche proprietà (coerenza spaziale, monocromaticità, brillanza), ha trovato importanti applicazioni in biologia e medicina fin dai primi mesi della sua esistenza. Col progredire della tecnologia laser (in particolare con l'avvento di laser accordabili in lunghezza d'onda e capaci di generare impulsi di luce ultracorti), tali applicazioni si sono ulteriormente estese. Oggi riguardano una larghissima varietà di impieghi cui si dedicano, in tutto il mondo, migliaia di ricercatori: fisici e biologi, chimici, medici e ingegneri impegnati in un comune sforzo di ricerca interdisciplinare.

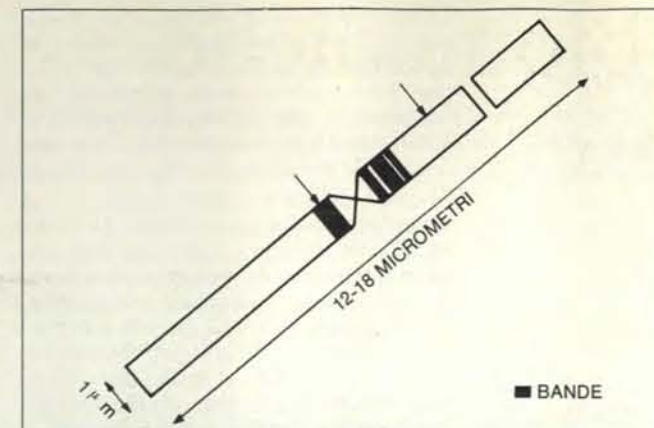
Le applicazioni biomediche del laser si possono distinguere in due grandi categorie: quelle in cui la luce viene impiegata per studiare le proprietà del campione biologico, senza alterarle, e quelle in cui la luce laser altera sensibilmente tali proprietà. Naturalmente tutte si basano su particolari aspetti dell'interazione tra la radiazione laser e il campione biologico: diffusione (elastica o anelastica) della luce, assorbimento, emissione fluorescente. Nella grande varietà di fenomeni studiati, tali effetti possono verificarsi separatamente oppure mescolati in varie misure; presentarsi da soli oppure accompagnati da complessi fenomeni biofisici e biochimici; verificarsi su campioni macroscopici o, concentrando il fascio ai limiti consentiti dalla diffrazione, su campioni microscopici subcellulari; avvenire in regime stazionario con fasci di illuminazione continui o in regime transitorio su scale di tempi molto brevi impiegando impulsi di eccitazione ultracorti.

La descrizione che segue non può che limitarsi ai principali esempi di tali applicazioni e si propone di illustrarne soprattutto i principi e le potenzialità.

Le applicazioni biologiche del laser finora studiate appartengono, in grande maggioranza, alla prima delle due categorie precedentemente menzionate. La coerenza spaziale viene frequentemente utilizzata per concentrare il fascio su singole cellule o su organelli cellulari al fine di studiarne le proprietà tramite l'analisi della luce riemessa per fluorescenza o diffusa. La monocromaticità consente di ottenere selettività nell'eccitazione di composti fluorescenti e accuratezza nell'analisi spettrale della luce diffusa. L'impiego di impulsi ultracorti, infine, consente di studiare fenomeni molto rapidi.

Tra le applicazioni appartenenti alla seconda categoria ci limitiamo a citare la «microchirurgia cellulare», descritta qui di seguito. In essa il fascio focalizzato possiede un'intensità sufficiente ad alterare le proprietà del campione (di solito la parte di una cellula) irraggiato.

La possibilità di concentrare la luce laser su un'areola di dimensioni minime dell'ordine della lunghezza d'onda (tipicamente un micrometro) fu ben presto utilizzata dopo la realizzazione del primo laser (avvenuta nel 1960), in esperimenti di biologia cellulare. Risale infatti al 1962 il primo esperimento di irraggiamento di organelli cellulari con un laser a rubino, realizzato presso l'Ecole Pratique des Hautes Etudes di Parigi. A partire da quella data, la tecnica di irraggiare selettivamente con un «microfascio» laser diverse strutture cellulari venne sviluppata in diversi laboratori acquistando una crescente importanza per la ricchezza dei risultati forniti. Lo scopo di questa tecnica è duplice: da un lato studiare i meccanismi di interazione tra la radiazione laser e diversi sistemi cellulari e dall'altro studiare il comportamento cellulare conseguente all'alterazione di specifici organelli o regioni cellulari e da ciò valutarne le funzioni e la relativa importanza. La strumentazione più avanzata per questa applicazione si trova probabilmente presso il Department of Developmental and Cell Biology dell'Università della California, a Irvine. Tale strumentazione si avvale di diversi tipi di laser (ad argo, a Nd-YAG, cioè neodimio - yttrium aluminum garnet, a coloranti), operanti sia in continua sia a impulsi, e di numerose armoniche ottiche delle frequenze emesse. Il fascio viene inviato sul campione tramite uno speciale microscopio, la cui ottica ne consente l'osservazione durante l'irraggiamento, tramite una camera per ripresa televisiva, e un monitor. Lo strumento, che possiede un potere risolutivo spaziale di circa 0,25 micrometri, consente di eseguire raffinate operazioni di chirurgia cellulare. Il trattamento del campione con sostanze fotoattivabili (ad esempio coloranti acridinici e psoraleni), che si legano in modo specifico a diversi costituenti cellulari, amplia considerevolmente le possibilità di indagine. Ad esempio, il trattamento con psoraleni e l'irraggiamento con la riga a 365 nanometri del laser ad argo consente di alterare selettivamente il DNA e/o l'RNA nella cromatina e negli organelli che contengono acidi nucleici. Tra le molte ricerche in corso (studi di struttura e funzioni di cromosomi, di mitosi, di funzionamento e ultrastruttura cellulare ecc.) desideriamo ricordare uno studio della contrattilità delle cellule del miocardio eseguito nel laboratorio citato. Focalizzando il fascio laser sui singoli mitocondri di cellule del ventricolo cardiaco è stato possibile alterare in modo controllato la figura di contrazione di singole cellule (e di quelle elettricamente accoppiate alla cellula irraggiata) e correlarla con le variazioni del potenziale di azione della membrana e verificare, inoltre, che l'effetto, possibile solo nelle cellule con potenziale d'azione di tipo «pace-maker», è connesso a una piccola alterazione della membrana cellulare direttamente al di sopra del mitocondrio irraggiato. Tale ricerca è di grande importanza per la comprensione del funzionamento del muscolo cardiaco.

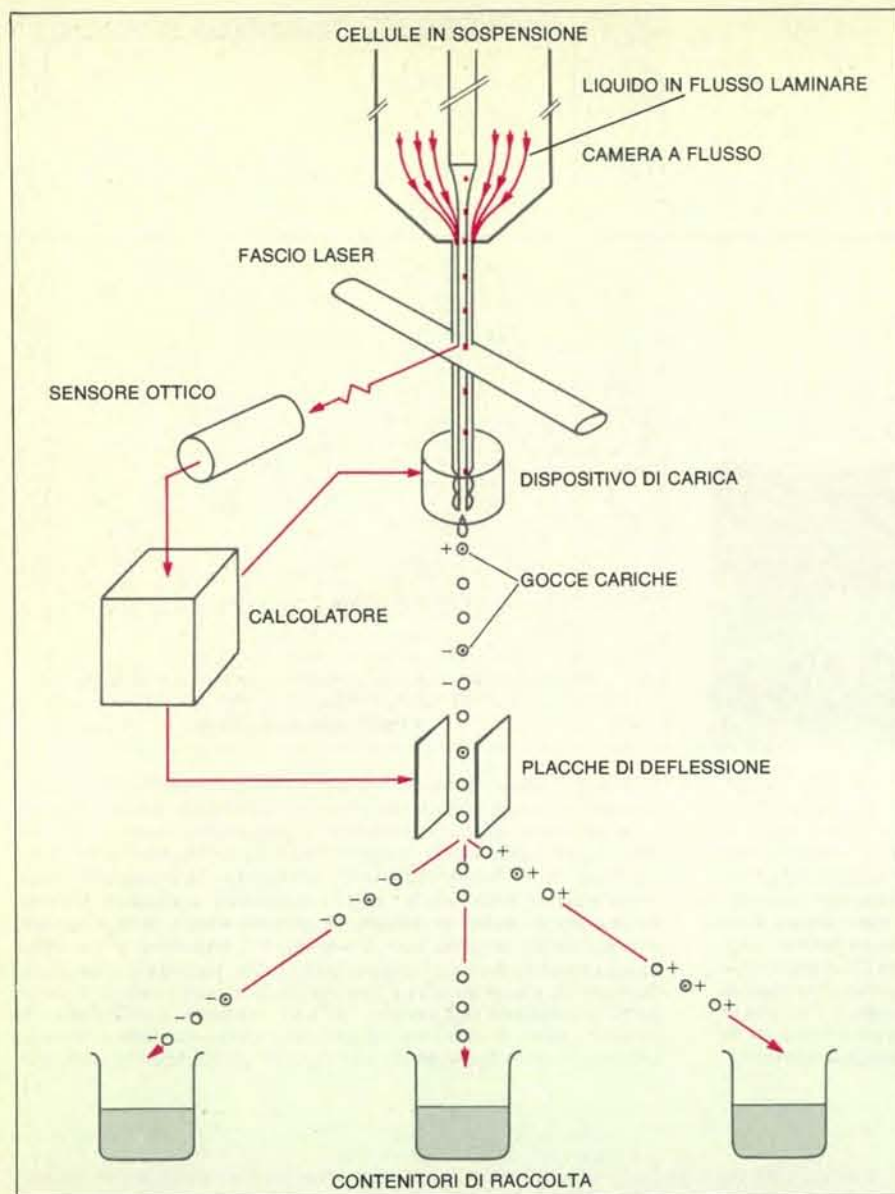
L'uso di una strumentazione concettualmente analoga (comprendente un microscopio), ma con intensità di irraggiamento sostanzialmente più basse (e tipicamente non superiori a 10^6 watt per centimetro quadrato), consente di studiare la fluorescenza emessa da campioni biologici microscopici.

I cromosomi di ogni specie animale o vegetale, trattati con coloranti acridinici e illuminati con luce ultravioletta, presentano caratteristiche figure di emissione fluorescente. Tali figure, sottili bande brillanti (circa un micrometro) su un fondo uniforme, sono utilizzate per il riconoscimento e la classificazione dei cromosomi umani; l'alterazione di tale figura nei cromosomi umani può essere correlata a particolari situazioni patologiche. L'origine delle figure non è ancora chiara: o sono dovute a un maggior contenuto di molecole fluorescenti (ad esempio per una maggiore accessibilità del DNA) oppure alla presenza di un DNA più ricco in sequenze adenina-timina che provocano una maggiore efficienza di emissione da parte delle molecole di colorante intercalate. Per verificare l'ipotesi più corretta è stato compiuto un esperimento sul cromosoma M della fava *(Vicia faba)*, trattato con il colorante mostarda di chinacrina.

Un impulso di luce laser della durata di 0,4 nanosecondi è stato focalizzato, con dimensioni inferiori alla larghezza delle bande, successivamente nelle due regioni indicate dalle frecce. Il segnale di fluorescenza ottenuto nella regione della banda (rappresentato dai punti) è stato ridotto in ampiezza di un fattore 1,54 per confrontarne la forma con l'altro (rappresentato dalla linea intera). La sostanziale uguaglianza di forma tra due impulsi indica la validità, nel caso esaminato, della prima ipotesi. Infatti nel secondo caso, i due impulsi avrebbero avuto forme sostanzialmente diverse. L'esperimento è stato eseguito con un microfluorometro a laser impulsato con alta risoluzione spaziale (0,3 micrometri) e temporale (0,2 nanosecondi). Lo strumento è stato realizzato presso il Centro di elettronica quantistica e strumentazione elettronica nell'ambito del Progetto finalizzato «Tecnologie biomediche» del CNR.

L'analisi dell'emissione fluorescente è molto utile per lo studio di proprietà strutturali di biomolecole e di proprietà funzionali di cellule. L'emissione può essere dovuta alla biomolecola stessa (fluorescenza primaria), generalmente eccitabile nell'ultravioletto (come per i costituenti del DNA), oppure a molecole di sostanze coloranti (traccianti fluorescenti) che si legano a una specifica posizione della biomolecola (fluorescenza secondaria). In questo caso può risultare particolarmente utile la misura del tempo di decadimento della fluorescenza, un parametro assai sensibile nell'interazione tra biomolecola e colorante. Per questa misura l'uso di impulsi laser di durata inferiore al nanosecondo (10^{-9} secondi, limite ottenibile con lampade convenzionali), risulta particolarmente adatto. Inoltre, se il laser è accordabile, la lunghezza d'onda può essere «sintonizzata» sul picco di assorbimento della sostanza fluorescente, ottenendo una selettività di eccitazione spettrale, oltre che spaziale, su singole areole cellulari.

Tra le molte ricerche in corso in questo campo (riguardanti le reazioni enzimatiche con emissione fluorescente, la relazione tra potenziali di membrana e l'assunzione cellulare di sostanze fluorescenti, la struttura secondaria di acidi nucleici) desideriamo ricordare lo studio delle proprietà di fluorescenza di complessi formati da coloranti acridinici col DNA. Questa ricerca è stata compiuta in collaborazione tra due centri del CNR, quello di Elettronica quantistica e strumentazione elettronica di Milano e quello di Istochimica di Pavia.

È noto che i coloranti acridinici formano un complesso col DNA in cui la molecola di colorante si intercala tra due coppie di basi adiacenti della doppia elica del DNA. Le proprietà di fluorescenza di tali complessi molecolari sono così sintetizzabili: il tempo di decadimento della fluorescenza (e perciò il rendimento quantico dell'emissione) viene considerevolmente aumentato (fino a qualche decina di nanosecondi) rispetto a quello della molecola di colorante (qualche nanosecondo) se questa intercala una sequenza di due coppie di basi adenina-timina (A-T). Al contrario tale tempo diminuisce nettamente (fino a qualche decimo di nanosecondo) se almeno una delle coppie di basi è del tipo guanina-citosina (G-C). Ne consegue una possibilità di indagine sulla sequenza di coppie di basi del DNA tramite l'analisi dell'emissione fluorescente previa «colorazione» con un colorante acridinico. Tale analisi non è semplice in quanto le fluorescenze dei due tipi di complessi non si sovrappongono linearmente. Infatti i due complessi, possono interagire, se distano tra loro di non più di alcuni nanometri, secondo un modello in cui i complessi del primo tipo («donatori») trasferiscono la loro energia di eccitazione a quelli del secondo («accettori»).

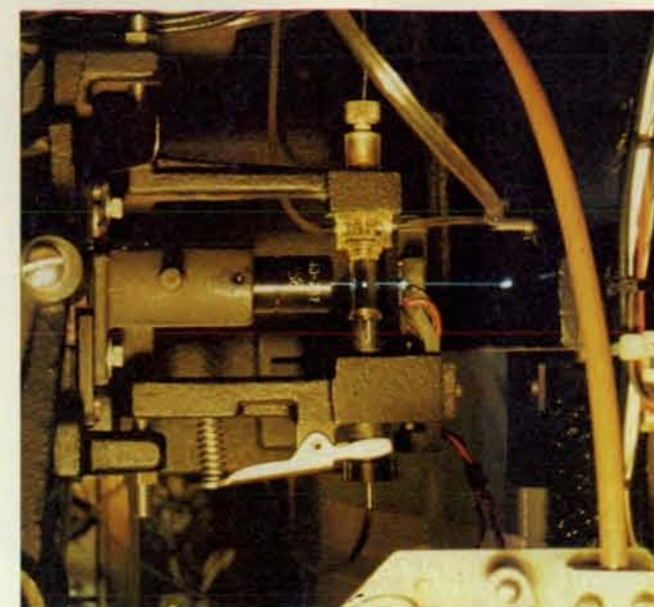
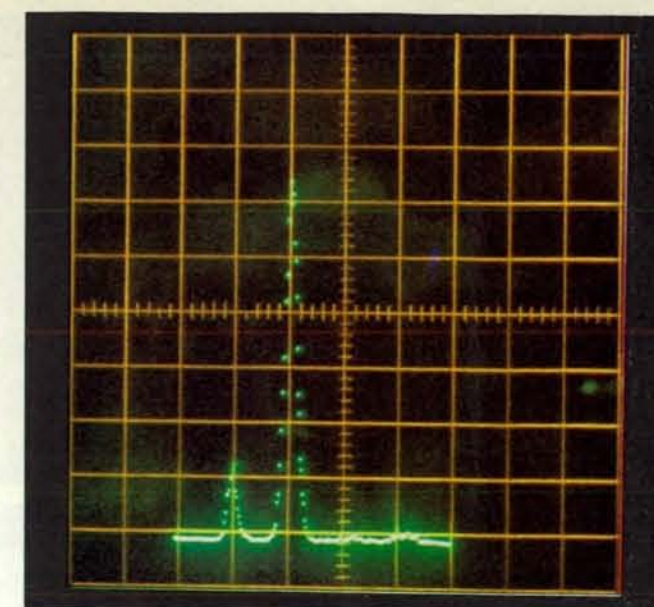
Le ricerche compiute presso i due centri suddetti hanno consentito, sulla base delle proprietà dei complessi citati, di caratterizzare lo stato funzionale della cromatina nelle fasi del ciclo cellulare di alcuni tipi di cellule, tramite il tempo di decadimento della fluorescenza e di compiere esperimenti interpretativi sulla bandatura fluorescente di cromosomi trattati con coloranti acridinici. In questo secondo caso la bandatura si evidenzia trattando i cromosomi anche con alcuni tipici farmaci antitumorali (la daunomicina e l'attinomicina D) che sembrano legarsi al DNA per intercalazione come i coloranti acridinici menzionati.

Rappresentazione schematica di un sistema per citometria a flusso. Nel sistema *slit scan flow*, il fascio laser è focalizzato lungo una linea sottile e consente quindi una mappatura della fluorescenza, lungo la direzione del flusso, della cellula che lo attraversa. Ciò consente, ad esempio, di eseguire un confronto tra la fluorescenza del nucleo cellulare e quella del citoplasma utilizzato in tecniche di prescreening per la diagnosi del cancro cervicale. Operando con cromosomi umani che attraversano il fascio secondo la loro lunghezza (da 2 a 20 micrometri), trattati con un colorante specifico per il DNA, si ottiene, tramite l'osservazione della fluorescenza, il profilo spaziale della distribuzione del DNA, anziché il suo contenuto totale come nella citometria ordinaria. Ciò dovrebbe consentire di risolvere tutti i tipi di cromosomi umani e di rilevare eventuali anomalie cromosomiche. Cellule o, in particolare cromosomi, che forniscono un segnale di fluorescenza entro soglie prefissate possono essere separate dal resto della sospensione e raccolte, provocando la suddivisione in gocce del flusso uscente dalla camera con un trasduttore piezoelettrico, caricando elettricamente le gocce e deflettendole tramite un campo elettrico comandato dal segnale di fluorescenza.

È opportuno ancora ricordare che l'analisi del segnale di fluorescenza eccitato da un laser a impulsi in un'areola cellulare consente di studiare la distribuzione spaziale, all'interno di una singola cellula, di farmaci fluorescenti (quali il chemioterapico adriamicina) e la loro interazione coi costituenti cellulari tramite l'analisi della forma d'onda di fluorescenza.

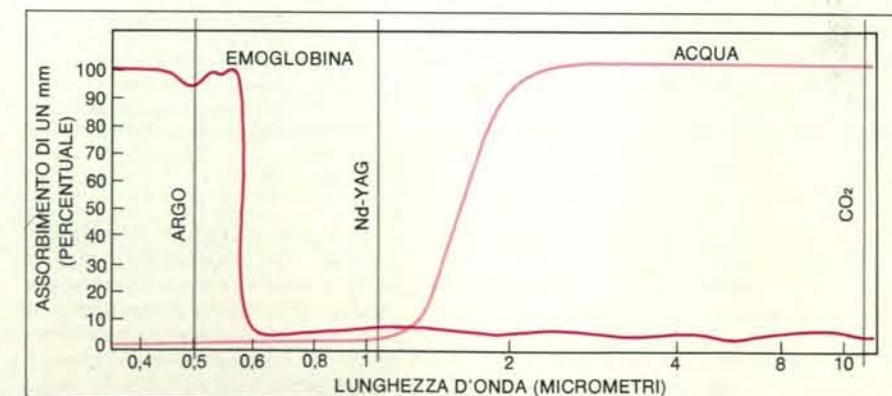
Un'altra applicazione basata principalmente sull'osservazione della fluorescenza emessa da un numero molto elevato di cellule osservate singolarmente, previo trattamento con marcanti fluorescenti, si ha nella «citometria a flusso». In questa tecnica le cellule, in sospensione, sono poste in una speciale camera a flusso ove vengono allineate per attraversare, una a una, un intenso e focalizzato fascio di luce laser. Ogni cellula, attraversando il fascio, produce uno o più segnali ottici (di fluorescenza e di diffusione), che vengono rivelati e analizzati. Il laser più usato in questa tecnica è quello ad argo la cui riga blu a 488 nanometri è particolarmente adatta all'eccitazione della fluorescenza di molte sostanze che si legano in modo specifico a diversi costituenti cellulari. La tecnica consente, ad esempio, di valutare, in ogni cellula, il contenuto di DNA, di RNA, o di proteine. Analogamente, la presenza di anticorpi fluorescenti consente di studiare la presenza di virus in colture cellulari infette. All'analisi della fluorescenza si può aggiungere lo studio della luce diffusa in varie direzioni, e ciò consente di acquisire informazioni sulle dimensioni e sulla morfologia cellulare. Recentemente la tecnica è stata raffinata focalizzando il fascio laser lungo una striscia sottile (larga, ad esempio, un micrometro e lunga qualche decina) situata trasversalmente alla direzione del flusso. Ciò consente di eseguire una scansione lungo ogni singola cellula *(slit-scan flow system)* ottenendo una mappatura cellulare dei costituenti in esame.

La citometria a flusso è l'unica tecnica che fornisca misure quantitative su un numero di cellule molto elevato in modo molto rapido (fino a 5×10^4 cellule al minuto), consentendo così un'alta precisione statistica. Diverse versioni commerciali di questi strumenti esistono sul mercato da alcuni anni e vengono correntemente impiegate in numerosi ospedali e centri di ricerca per analisi dei cicli cellulari, diagnosi di tumori, studio di malattie genetiche, di effetti di farmaci ecc.

La diffusione della luce può essere utilizzata in vari modi per acquisire informazioni su un campione biologico. Come già accennato, la diffusione elastica (cioè senza variazioni di frequenza) da parte di cellule dà indicazioni sulle loro dimensioni. La diffusione anelastica prodotta per effetto Raman da molecole biologiche (con spostamenti di frequenza tipicamente superiori a 100 cm^{-1}) dà informazioni sulle vibrazioni interne alle molecole stesse (la cui frequenza è pari allo spostamento in frequenza della luce osservato). La diffusione «quasi-elastica», caratterizzata da spostamenti di frequenza minori, dà a sua volta informazioni su moti più lenti generalmente connessi a interazioni intermolecolari: viene impiegata, ad esempio, per lo studio della formazione di micelle, importanti in molti fenomeni biologici.

Lo spostamento in frequenza della luce nella diffusione quasi-elastica è interpretabile come un effetto Doppler dovuto alla velocità degli oggetti diffondenti. Se questi possiedono un moto browniano, la luce diffusa presenta un allargamento spettrale attorno alla frequenza incidente. Se invece possiedono un moto di trascinamento unidirezionale (come ad esempio i globuli rossi nella circolazione sanguigna), la luce diffusa presenta uno spostamento in frequenza. La misura di tale spostamento consente una determinazione «remota» e precisa della velocità dell'oggetto diffondente e la tecnica prende il nome di velocimetria laser-Doppler. L'alta frequenza e la monocromaticità della luce laser consentono di spingere i limiti della misura a velocità minime di pochi micrometri al secondo.

La fotografia a sinistra illustra un particolare del sistema per citometria a flusso in funzione presso il Centro di istochimica del CNR all'Istituto di anatomia comparata dell'Università di Pavia. La curva *(a destra)* fornisce il numero di cellule (in ordinate) in funzione del contenuto di DNA di ogni cellula (in ascisse). Le cellule, trattate con ioduro di propidio, sono state ottenute da un prelievo bioptico osseo di un paziente affetto da leucosi. I due picchi si riferiscono a due popolazioni cellulari in fase G_1. In un campione normale sarebbe presente un solo picco.

La coerenza spaziale del laser consente, come nelle applicazioni già descritte, di eseguire le misure, con un fascio concentrato, su campioni microscopici.

Tra le molte applicazioni biomediche di questa tecnica ci limitiamo a ricordare, in primo luogo, lo studio dell'idrodinamica intracellulare tramite l'osservazione di correnti di materiale protoplasmatico (sostanze nutritive e materiale genetico) all'interno di singole cellule; quest'applicazione rappresenta il più fine livello di dettaglio raggiungibile nella velocimetria di sistemi biologici. In secondo luogo va ricordato lo studio della motilità di spermatozoi (e, in generale di microrganismi) in sospensione. La motilità degli spermatozoi, che ovviamente è parte importante della loro funzione, provoca un sostanziale allargamento spettrale della luce diffusa, rispetto a campioni privi di motilità. La tecnica, pur inferiore all'osservazione (anche cinematografica) al microscopio per l'analisi dei dettagli del moto, è assai utile per la rapidità nel fornire indicazioni di massima. Infine, per quanto riguarda lo studio del flusso sanguigno, la tecnica consente di studiare sia la microcircolazione sottocutanea, tramite l'analisi spettrale della luce diffusa dalla pelle, sia la circolazione in grandi vasi, tramite l'inserimento di un piccolo catetere a fibra ottica che raccoglie anche la luce retrodiffusa da analizzare. Un'applicazione particolarmente importante (la prima in campo medico, iniziata nel 1972) si ha nello studio del flusso sanguigno nel sistema vascolare della retina, interessato da malattie sia specifiche sia sistemiche come diabete e ipertensione. Rispetto alle tecniche convenzionali la velocimetria laser-Doppler presenta il vantaggio di essere rapida, sicura e «non invasiva».

Il laser, nel regime di funzionamento di *mode-locking* (agganciamento di fase tra i modi oscillanti), consente di generare impulsi di luce ultracorti. Numerosi laboratori possiedono laser capaci di generare impulsi della durata di qualche picosecondo (10^{-12} secondi). In altri laboratori, più specializzati, si è raggiunta una durata minima inferiore a 50 femtosecondi (50×10^{-15} secondi). La disponibilità di tali impulsi, non ottenibili con altre tecniche, ha aperto la possibilità di iniziare un'ampia gamma di ricerche sperimentali riguardanti fenomeni ultrarapidi. Tra

Rappresentazione schematica dello spettro di assorbimento dell'emoglobina e dell'acqua; le linee verticali rappresentano le righe di emissione dei principali laser di interesse fototerapeutico: laser ad argo (riga blu a 0,488 micrometri e riga verde a 0,5145 micrometri), laser a Nd-YAG (lunghezza d'onda 1,06 micrometri) e laser a CO_2 (lunghezza d'onda 10,6 micrometri). L'assorbimento da parte dell'emoglobina *(curva in colore intenso)* risulta assai elevato per il laser ad argo e, in generale, per tutti i laser che emettono tra 0,4 e 0,6 micrometri. Nell'irraggiamento di un vaso sanguinante l'assorbimento di luce da parte dell'emoglobina provoca la rottura dei globuli rossi, l'attivazione delle piastrine e danni alle pareti interne del vaso stesso. Questa sequenza di eventi porta alla coagulazione del sangue e pone termine al sanguinamento del vaso, effetto che rende i laser che emettono nel visibile particolarmente adatti alla coagulazione di emorragie superficiali. L'irraggiamento di un tessuto con un laser a CO_2 di intensità sufficiente ne provoca l'incisione tramite la vaporizzazione dell'acqua contenuta (80-90 per cento). A causa del forte assorbimento, la penetrazione del fascio di un laser a CO_2 nei tessuti è limitata a circa 50 micrometri. Perciò questo laser è particolarmente adatto alla vaporizzazione e al taglio dei tessuti. La regione spettrale compresa tra 0,6 e 1,5 micrometri, ove l'assorbimento dell'emoglobina e dell'acqua *(curva in colore chiaro)* è basso, consente alla radiazione di penetrare nei tessuti per circa un millimetro. Il laser a Nd-YAG è per questo motivo particolarmente adatto alla coagulazione in caso di emorragie massive.

La fotografia in alto mostra un sistema per la generazione di impulsi di luce ultracorti in dotazione al Centro di elettronica quantistica e strumentazione elettronica del CNR al Politecnico di Milano. Un laser ad argo, parzialmente visibile a destra, viene fatto funzionare in regime di *mode-locking* (agganciamento di fase tra i modi oscillanti del risonatore laser) tramite un modulatore acustico-ottico, inserito nella cavità. Gli impulsi di luce (a una lunghezza d'onda di 514,5 nanometri) pompano in modo sincrono un laser a colorante (rodamina 6G), visibile al centro. Il fascio emesso da questo laser, costituito da un treno di impulsi della durata di qualche picosecondo alla lunghezza d'onda di 580 nanometri, viene diviso in due parti. Una *(al centro della fotografia)* viene inviata al sistema per la misura di durata degli impulsi, ottenuta con tecniche di generazione di seconda armonica ottica. L'altra viene inviata su una celletta contenente il campione di cui si vuole studiare la fluorescenza *(luce rossa a sinistra)*. Le altre due fotografie in basso mostrano i particolari del laser a colorante e della celletta del campione.

questi, i più importanti nel campo della biologia riguardano le fasi iniziali dei processi della fotosintesi e della visione.

Nella fotosintesi la luce viene assorbita da speciali pigmenti («*antenna pigments*») che in un tempo dell'ordine del picosecondo, tramite interazioni non radiative dipolo-dipolo, trasferiscono l'energia di eccitazione acquisita a molecole di clorofilla. Lo stato di eccitazione («eccitone») migra successivamente tra diverse molecole di clorofilla contigue finché si avvicina a un «centro di reazione» ove viene catturato e utilizzato per iniziare il processo fotosintetico. Tra l'assorbimento della luce da parte del pigmento e la cattura da parte del centro di reazione, pur intercorrendo diverse centinaia di singoli trasferimenti di energia, passa meno di un nanosecondo. Le varie tecniche di spettrometria a picosecondi possono utilmente essere impiegate per studiare il sistema d'antenna delle piante (pigmenti, clorofilla, centro di reazione). Ad esempio, si può eccitare il sistema con impulsi laser a picosecondi di opportuna lunghezza d'onda e osservare contemporaneamente l'emissione fluorescente di varie specie molecolari tramite una camera a scansione («*streak camera*») ultrarapida. La fluorescenza dà un'indicazione della eccitazione delle specie osservate. In particolare nelle piante verdi il decadimento della fluorescenza della clorofilla dipende dal tempo di migrazione eccitonica verso il centro di reazione. La sua misura è utile per la definizione del modello teorico che descrive il sistema.

La visione nei vertebrati sembra basarsi su un processo primario di isomerizzazione del cromoforo retinale della rodopsina presente nelle cellule fotosensibili (bastoncini) della retina. Tale processo comporterebbe la formazione di alcuni prodotti intermedi di cui il primo si ritiene sia la «prelumirodopsina». La cinetica di formazione di questo prodotto è stata studiata in soluzione, negli ultimi anni, impiegando impulsi di luce ultracorti (6 picosecondi di durata) generati da un laser a neodimio alla lunghezza d'onda di 1060 nanometri. Tali impulsi sono stati successivamente sottoposti a duplicazione di frequenza ottica, lunghezza d'onda di 530 nanometri, con un cristallo KDP (kalium dihydrogenate phosphate) ed effetto Raman stimolato in una cella di benzene per portare la lunghezza d'onda a un valore di 561 nanometri prossimo al picco di assorbimento della prelumirodopsina. L'impulso a 530 nanometri è stato impiegato per eccitare il processo e quello a 561 nanometri per «sondare», tramite il suo assorbimento, la creazione di prelumirodopsina. Al fine di seguire nel tempo questo processo, l'impulso a 561 nanometri è stato a sua volta suddiviso in un treno di 15 impulsi di sonda (tutti di 6 picosecondi), tramite un reticolo in riflessione a gradinata (echelon). Variando l'echelon è stato possibile variare il ritardo tra due impulsi successivi nell'intervallo 2-20 picosecondi. La misura ha indicato che l'assorbimento dell'impulso a 561 nanometri, attribuito alla prelumirodopsina, avviene già entro la durata del primo impulso di sonda. Dunque la formazione di questa sostanza deve avvenire entro 6 picosecondi dall'eccitazione. Questo risultato è stato assunto a conferma del modello che considera la prelumirodopsina come il primo composto formato dopo l'eccitazione della rodopsina e dà indicazioni sui livelli energetici della molecola coinvolti nel processo. Studi assai recenti, pure compiuti con tecniche laser a picosecondi, hanno tuttavia dimostrato che il processo comporta la formazione di altri prodotti intermedi e richiede ulteriori approfondite ricerche.

Le applicazioni mediche del laser si basano sulle proprietà di assorbimento della radiazione da parte di alcuni costituenti essenziali del corpo umano, quali l'acqua (presente nei tessuti in percentuale dell'80-90 per cento) e il sangue, e di alcuni pigmenti (come, ad esempio, la melanina). Naturalmente tali applicazioni, che comportano una sensibile alterazione del «campione» irraggiato, si avvalgono anche delle proprietà caratteristiche del laser come, di volta in volta, vedremo.

La prima applicazione medica del laser si ebbe nel 1962, con l'impiego di un laser a rubino per la saldatura della retina nell'occhio, successivamente diffusasi in tutto il mondo con l'impiego del laser ad argo. L'invenzione del laser a CO_2, la cui radiazione è fortemente assorbita dall'acqua, aprì, dopo il 1967, il vasto campo delle applicazioni chirurgiche.

La disponibilità di fibre ottiche, entro le quali solo la luce molto collimata del laser può essere inviata con alta efficienza, consentì, a partire dal 1974, di combinare le tecniche laser con quelle dell'endoscopia convenzionale. Fu così possibile iniziare il trattamento di emorragie interne col laser ad argo, la cui luce verde-blu è fortemente assorbita dall'emoglobina, cui si aggiunse poco più tardi, nel 1975, il laser a Nd-YAG la cui radiazione, più penetrante nei tessuti, ne consente un ampio impiego in endoscopia. Un altro campo di applicazione di più recente sviluppo («fotochemioterapia») utilizza l'assorbimento della luce da parte di sostanze opportunamente introdotte nell'organismo, cui fa seguito una serie di reazioni biochimiche che portano all'effetto terapeutico desiderato.

Come già ricordato, gli oculisti sono stati i primi, tra i medici, a introdurre il laser nel loro strumentario clinico. L'assorbimento della luce nel fondo dell'occhio da parte di pigmenti dell'epitelio (in particolare la melanina) e della coroide, unito alla possibilità di concentrare il fascio su una superficie di dimensioni microscopiche, consente di curare numerose vasculopatie retino-coroidali e i processi degenerativi che conducono al distacco della retina. In quest'ultimo caso, infatti, il calore prodotto dall'assorbimento dell'energia laser nell'epitelio e nella coroide diffonde verso la sovrastante retina e la «cuoce», saldandola all'epitelio.

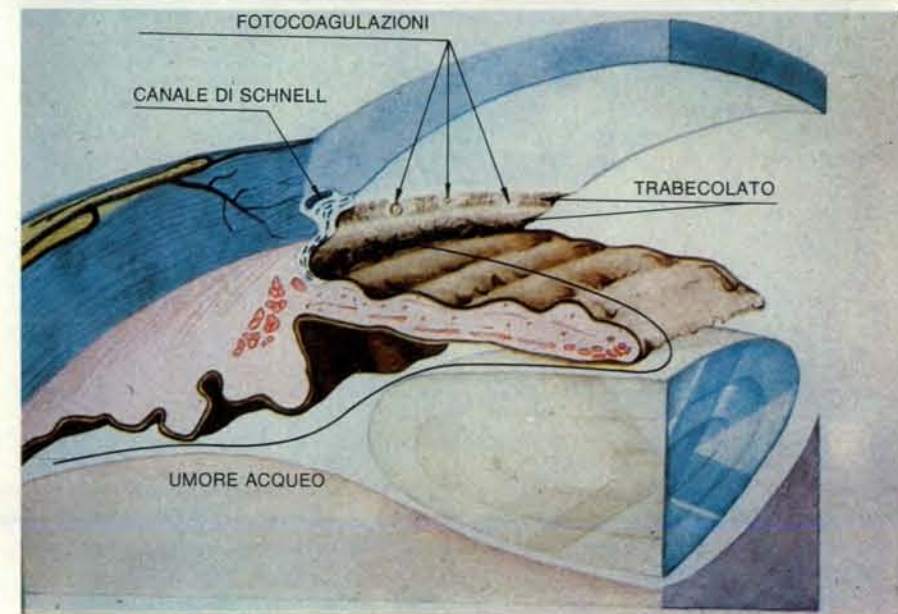
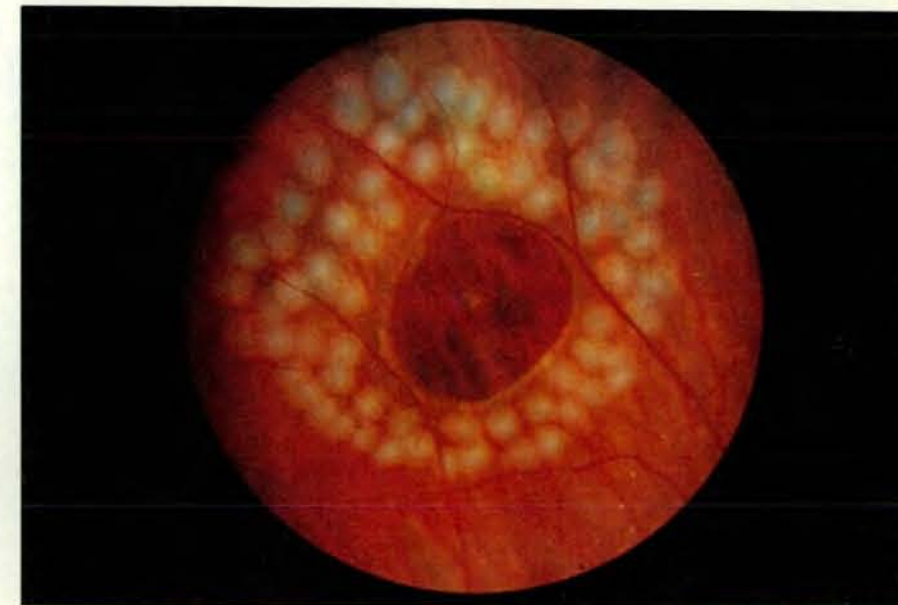
Il laser ad argo, universalmente impiegato in questa terapia, ha rappresentato un notevolissimo progresso rispetto al tradizionale fotocoagulatore a lampada ad arco allo xeno. Infatti, col laser, l'area irraggiata ha dimensioni minime; il tempo di esposizione è così breve da consentire di operare senza anestesia; i danni prodotti ai mezzi oculari dalla radiazione a largo spettro della lampada allo xeno, anche se filtrata, sono assenti.

Il laser ad argo risulta particolarmente utile anche nel trattamento delle retinopatie diabetiche. Si osservano in questa malattia sia la formazione di alterazioni microvascolari (dilatazioni, microaneurismi ecc.), che determinano una trasudazione di plasma nel tessuto retinico, sia la formazione di piccoli vasi (neovascolarizzazione retinica, dovuta all'ischemia retinica causata dalla malattia diabetica), che tendono facilmente a sanguinare. La distruzione con fotocoagulazione laser delle alterazioni microvascolari e delle aree retiniche ischemiche determina spesso una regressione delle lesioni retiniche e dei vasi neoformati patologici, rispettivamente.

La terapia del glaucoma rappresenta un'altra applicazione importante del laser. Nel caso del glaucoma ad angolo chiuso il deflusso dell'umore acqueo dalla camera anteriore dell'occhio alle strutture trabecolari è impedito dalla spinta in avanti della base dell'iride con chiusura dell'angolo irido-corneale. Si può in questo caso praticare col laser un foro nell'iride e consentire così un approfondimento della camera anteriore con l'apertura dell'angolo irido-corneale.

Anche i laser a stato solido a rubino e a Nd-YAG, nel funzionamento a impulsi giganti (durata di circa 20 nanosecondi e potenza fino a circa un megawatt), vengono utilmente impiegati in alcune patologie del segmento anteriore dell'occhio dalle scuole sovietica e svizzera. Presso quest'ultima scuola vengono compiuti, con il laser a Nd-YAG, interventi su membrane fibrose nel campo pupillare. Mentre l'intervento tradizionale prevede una discissione chirurgica, che risulta spesso assai delicata e difficile, un solo impulso laser consente di sezionare la membrana e aprire così completamente il foro pupillare.

L'invenzione del laser a CO_2 (1964) contribuì in modo decisivo alle applicazioni chirurgiche del laser che cominciarono tuttavia a diffondersi dopo il 1967, quando venne messo a punto il braccio articolato per il trasporto e la manipolazione del fascio.

Quando il fascio di un laser a CO_2 viene focalizzato su un tessuto, l'acqua che vi è contenuta in alta percentuale ne assorbe l'energia trasformandola in calore. Con un'intensità di pochi watt per millimetro quadrato l'acqua presente nel tessuto va rapidamente in ebollizione e il tessuto vaporizza. Se viene spostato lungo il tessuto, il fascio si comporta come un bisturi formando un'incisione la cui profondità dipende dall'intensità del fascio e dalla velocità di spostamento. La temperatura della zona incisa rimane costante, a 100 gradi centigradi, finché tutta l'acqua è evaporata e durante questo processo il calore diffonde

Esempi di applicazione laser in oftalmologia: la fotografia in alto mostra una rottura retinica circondata da fotocoagulazioni con laser ad argo disposte su più file. In basso viene mostrata schematicamente la struttura del segmento anteriore dell'occhio. Nel caso del glaucoma ad «angolo aperto», il deflusso regolare dell'umore acqueo è in parte impedito dall'appiattimento della struttura trabecolare. La tecnica di praticare dei fori nel trabecolato con laser a Nd-YAG o a rubino a impulsi «giganti» consente solo una ricanalizzazione temporanea del flusso, dato che dopo alcuni mesi i fori si richiudono. Al contrario, praticando delle fotocoagulazioni con il laser ad argo in continua, sulla fascia pigmentata del trabecolato, si ottiene una successiva retrazione cicatriziale, che allarga gli spazi trabecolari. Le frecce indicano i punti di tre fotocoagulazioni: la prima da sinistra è risultata eccessiva e ha formato una bollicina, le altre due sono corrette.

lateralmente verso i bordi del taglio. Ciò provoca un danno irreversibile ai tessuti per una profondità di 200-300 micrometri, ma provvede anche al sigillo dei piccoli vasi (di diametro inferiore a un millimetro) che confluiscono nella zona incisa. In assenza di vasi di calibro maggiore il taglio è privo di sanguinamento.

Accanto all'azione (relativamente) emostatica, di grande importanza nel trattamento di soggetti emofiliaci, e l'assenza di irraggiamento elettromagnetico utile nel trattamento di portatori di «pace-maker», altre importanti proprietà vengono attribuite a questo bisturi-laser: una azione batteriostatica, utile nel trattamento di parti infette e ustionate, una diminuzione del dolore post-operatorio e, soprattutto, una limitazione alla diffusione di cellule neoplastiche, per sigillo dei vasi ematici e linfatici, nella chirurgia dei tumori. Questi ultimi vantaggi non sono stati tuttavia ancora dimostrati e richiedono, per la loro verifica, accurati confronti con gruppi di controllo e *«follow-up»* su tempi lunghi. D'altro canto alcuni inconvenienti tecnici (ingombro, scarsa maneggevolezza, presenza di cavi e tubi di raffreddamento), la mancanza di contatto meccanico con il tessuto e l'alto costo hanno contribuito a limitare la diffusione di questo strumento nel campo della chirurgia generale.

L'abbinamento del laser a CO_2 con le tecniche di microchirurgia rappresenta il connubio più felice realizzato nel campo della chirurgia laser. Il fascio viene deflesso con precisione attorno all'asse di un microscopio operatorio tramite uno specchio comandato da micromanipolatori. L'ingrandimento fornito dal microscopio e l'accuratezza del posizionamento del fascio, individuabile con un fascetto di riferimento visibile, consentono di operare con grande precisione, in un'area limitata, utilizzando appieno i vantaggi offerti dall'impiego del laser. La microchirurgia laser si è affermata nei settori dell'otorinolaringoiatria, della ginecologia e, sia pure in modo meno ampio, della neurochirurgia.

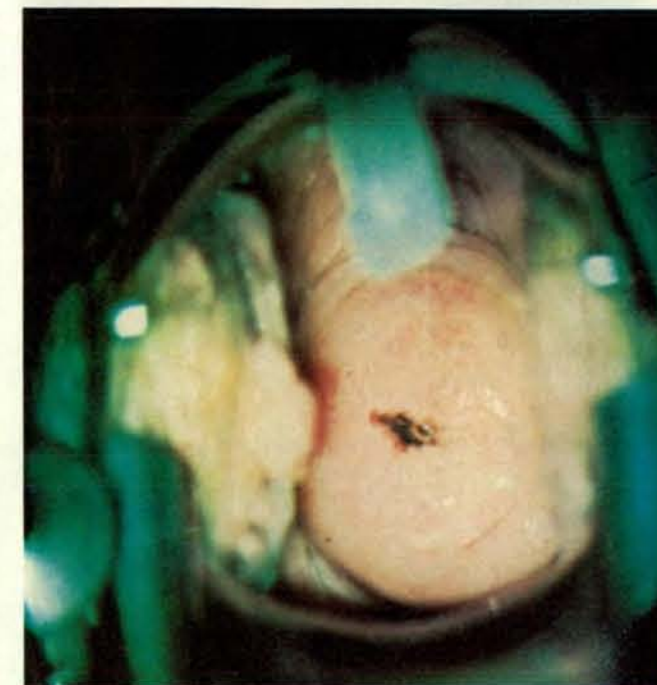
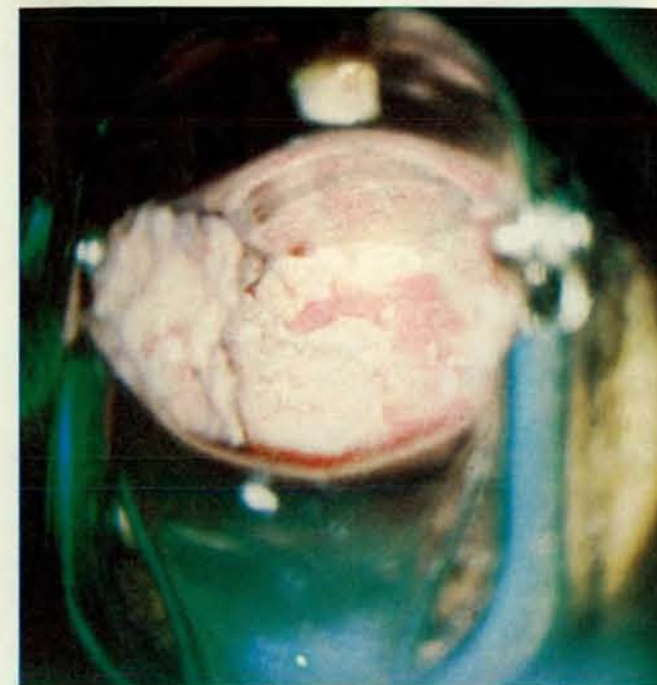
La principale indicazione, in otorinolaringoiatria, è la terapia delle lesioni endolaringee ove il chirurgo opera, a 40 centimetri di distanza, con un laringoscopio rigido, lungo e stretto. La totale assenza di sanguinamento (date le piccole dimensioni dei vasi) e l'impiego di un bisturi «immateriale» forniscono al chirurgo un completo controllo visivo del campo operatorio. Ciò consente un'accurata valutazione della profondità di incisione, tramite anche il preciso dosaggio dell'energia depositata senza contatto sul tessuto, il controllo del danno prodotto e il massimo rispetto dei tessuti sani. I risultati immediati sono decisamente superiori a quelli conseguibili con i metodi tradizionali: l'assenza di edema post-operatorio consente di evitare quasi sempre la tracheotomia e di ridurre sensibilmente la degenza ospedaliera. In alcuni centri questa chirurgia viene eseguita nell'ambito del «day hospital», e la maggioranza dei pazienti viene dimessa in giornata, anche dopo gli interventi di un certo impegno (cordectomia). Per quanto riguarda la patologia flogistica (polipi, angiomatosi delle corde vocali ecc.) anche i risultati a distanza sono ottimi: la possibilità di vaporizzare e asportare le lesioni senza ledere la muscolatura consentono un'ottima ripresa della funzionalità vocale.

Anche nel campo della microchirurgia ginecologica il laser a CO_2 consente di ottenere risultati migliori che con ogni altra tecnica. L'abbinamento del laser al colposcopio e al micromanipolatore consente un trattamento assai preciso in profondità e ampiezza, e l'intervento può essere fatto ambulatoriamente e, spesso, senza anestesia. La terapia laser risulta particolarmente indicata nelle lesioni precancerose della cervice. Il laser, in questo caso, viene usato come strumento di vaporizzazione del tessuto cervicale: è possibile raggiungere profondità di distruzione di 6-7 millimetri senza provocare sensibile dolore, con una zona di danno termico ai bordi dell'area trattata dieci volte meno estesa rispetto ai trattamenti convenzionali. Ne seguono una pronta guarigione locale, una ridotta fibrosi e assenza di stenosi, e, quindi, tra l'altro, la conservazione delle funzioni riproduttive e sessuali. Quando la lesione cervicale si estende nel canale cervicale è necessario ricorrere alla resezione della zona patologica (conizzazione). La conizzazione con laser a CO_2 offre vantaggi rispetto alla metodica tradizionale, quali la possibilità di effettuare gli interventi ambulatoriamente, con conseguente riduzione dei costi e delle complicazioni, quali il sanguinamento, e soprattutto permette la conservazione dell'integrità anatomica e funzionale, in quanto non si verificano fenomeni di stenosi del canale, né conseguente riduzione della fertilità. Anche per vari tipi di lesioni vulvari le metodiche chirurgiche con laser a CO_2 risultano particolarmente efficaci.

Negli interventi sul cervello e sul midollo spinale la precisione riveste la massima importanza al fine di non ledere centri e funzioni vitali. Per la stessa ragione, la manipolazione meccanica e le emorragie, che diminuiscono il controllo visivo, vanno ridotte al minimo. Anche in questo campo l'uso del laser a CO_2 connesso al microscopio operatorio ha consentito di vaporizzare con precisione grossi tumori e di eseguire con successo interventi diversamente irrealizzabili per la difficoltà di accedere alle lesioni da trattare.

Nel laser a CO_2 per uso chirurgico, il tubo laser è situato nella scatola a parallelepipedo disposta obliquamente sulla base mobile. Poiché le fibre ottiche trasparenti alla radiazione di 10,6 micrometri si trovano ancora in una fase iniziale di sviluppo, il trasporto del fascio dalla sorgente al campo operatorio viene realizzato mediante un braccio articolato (con gruppi di tre specchi agli snodi), terminante con un manipolo che possiede un puntale in cui è inserita una lente di seleniuro di zinco. Questa permette il passaggio della radiazione infrarossa del fascio laser e di quella visibile del fascio guida, emessa da un piccolo laser Ne-He. Il puntale con la lente di focalizzazione *(nel riquadro in alto a sinistra)* consente di usare il laser a CO_2 per chirurgia di superficie o endocavitaria a cielo aperto. Svitando il puntale, il braccio articolato può essere collegato direttamente al microscopio operatorio; la posizione del fascio laser viene ora regolata tramite un sensibile micromanipolatore, sempre con l'aiuto del fascio guida, nel campo del microscopio operatorio.

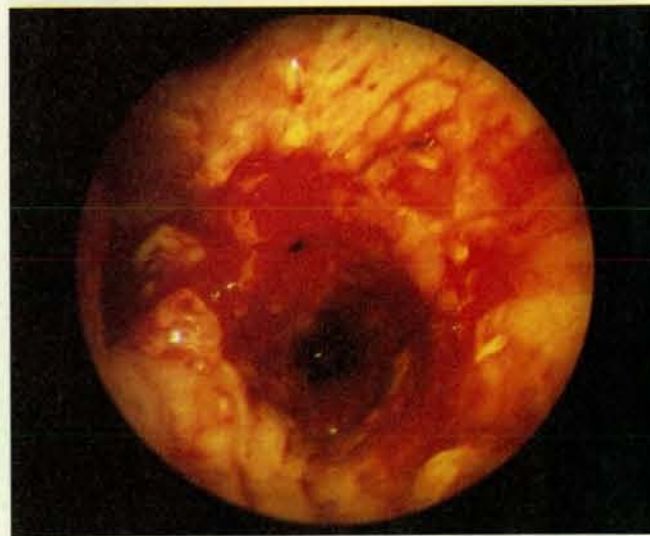
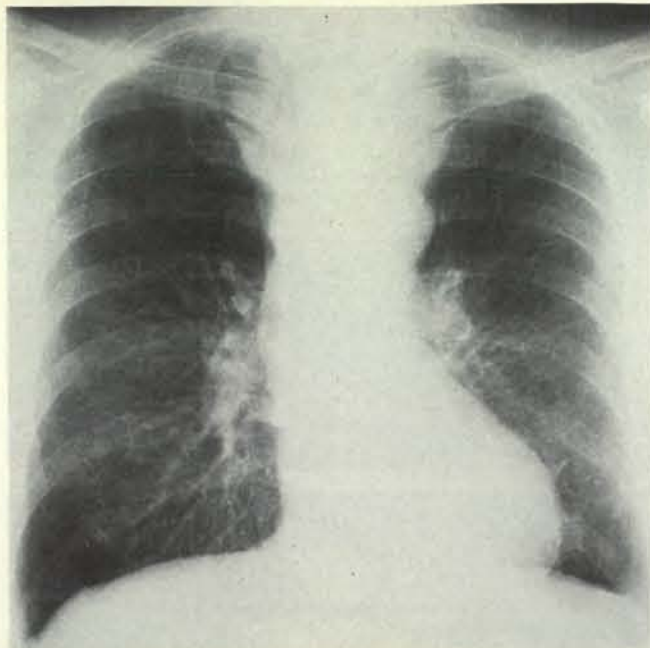
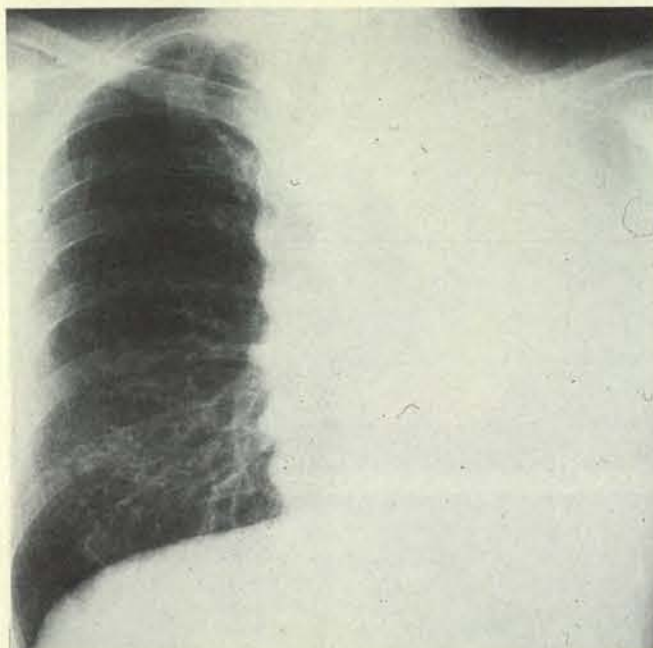
La fototerapia endoscopica rappresenta una delle più peculiari applicazioni mediche del laser. Essa si basa sulla possibilità di inviare la luce laser all'interno delle cavità del corpo umano (tratto gastrointestinale, bronchi, vescica) tramite una fibra ottica (sottile e flessibile guida di luce) inserita in un endoscopio convenzionale.

In gastroenterologia i laser ad argo e a Nd-YAG, la cui radiazione può essere trasportata da fibre ottiche, vengono impiegati con successo per il controllo di emor-

La serie di fotografie illustra due interventi clinici di microchirurgia con laser a CO_2. La sequenza in alto mostra un carcinoma delle corde vocali, prima e dopo la vaporizzazione con il laser. La rete vascolare è fitta, ma di piccolo calibro; l'intervento avviene perciò in ottime condizioni di emostasi. L'assenza di edema post-operatorio e la limitazione del danno ai tessuti sani permettono una rapida guarigione e un'ottima ripresa della funzionalità vocale. La sequenza in basso illustra la vaporizzazione di un condiloma acuminato della cervice uterina. La microchirurgia in campo otorinolaringoiatrico e ginecologico rappresenta la principale applicazione del laser a CO_2; essa viene praticata in vari centri e anche a livello ambulatoriale.

ragie. Il laser ad argo, la cui radiazione è fortemente assorbita dall'emoglobina (per il 63 per cento dopo 0,2 millimetri di sangue) viene impiegato, con potenze fino a 10 watt, per il trattamento di emorragie in piccoli vasi, o in regioni ricche di capillari. Il laser a Nd-YAG, assai meno assorbito sia dall'emoglobina sia dall'acqua, penetra per circa un millimetro nella mucosa ed è più adatto al controllo di emorragie massive o prodotte da vasi di maggior calibro. Questa maggior penetrazione e le potenze impiegate (fino a circa 100 watt) richiedono tuttavia molta cautela, per il rischio di perforazione della parete irraggiata. I casi più trattati si riferiscono a ulcere sanguinanti dello stomaco e del duodeno: alcune migliaia col laser a Nd-YAG (esito positivo nel 90 per cento e perforazione nel 2 per cento dei casi) e diverse centinaia con il laser ad argo (esito positivo nel 78 per cento, nessuna perforazione). Questi risultati dimostrano la grande efficacia della fotocoagulazione laser.

Studi comparativi rispetto alle terapie convenzionali con gruppi di pazienti di controllo e «randomizzazione», per quanto desiderabili, sono stati tuttavia compiuti solo recentemente e su poche decine di casi. Tali studi, infatti, risultano di difficile esecuzione sia per problemi di standardizzazione dei casi sia per ragioni etiche, in quanto coinvolgono spesso pazienti ad alto rischio. I risultati finora ottenuti indicano che il laser ad argo non fornisce differenze significative rispetto ai controlli, mentre il laser a Nd-YAG ha fornito significativi vantaggi sia nel controllo iniziale delle emorragie (particolarmente utile nella terapia d'urgenza), sia nella successiva riapertura delle lesioni trattate.

Applicazioni recenti stanno dimostrando che il laser a Nd-YAG, per la sua maggior profondità di penetrazione nei tessuti, consente di eseguire per via endoscopica anche interventi di distruzione di tumori o di rimozione di polipi. Particolarmente importante è la possibilità di ripristinare la canalizzazione in caso di neoplasie stenosanti dell'esofago e del grosso intestino, del retto e dei bronchi. Questi interventi, anche se di tipo palliativo, consentono di migliorare le condizioni del paziente e di recuperarlo per altre terapie.

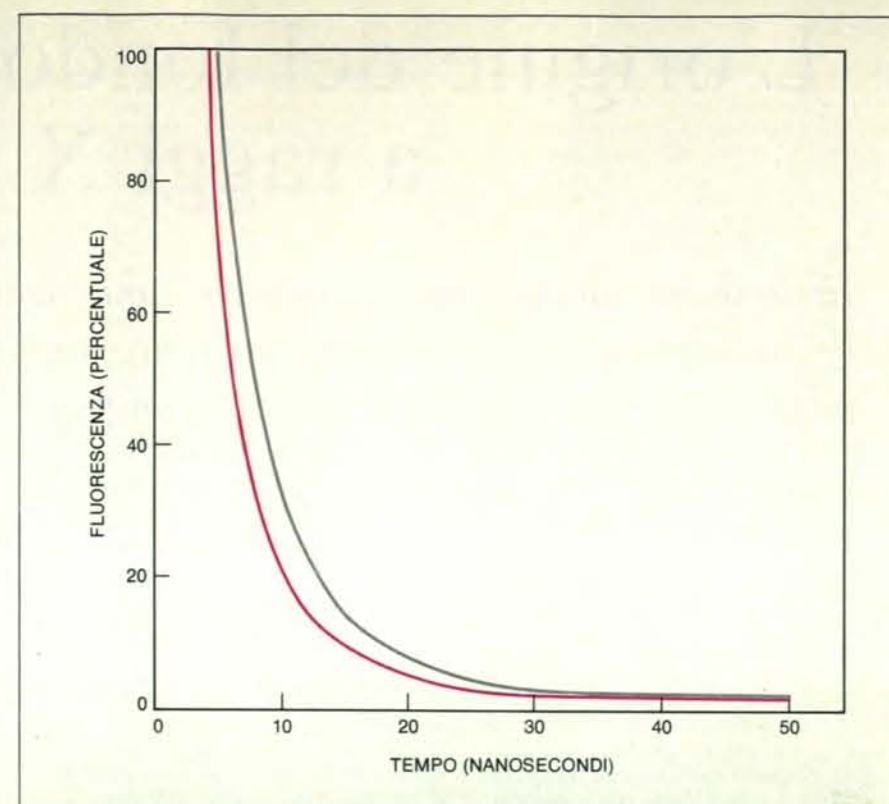
Per quanto riguarda le applicazioni endoscopiche in campo urologico le ricerche cliniche sono ancora in fase iniziale. I tumori della mucosa vescicale, a basso grado di infiltrazione, sono facilmente aggredibili col laser a Nd-YAG. L'irraggiamento laser provoca un istantaneo sbiancamento e rigonfiamento delle zone

La serie di fotografie mostra un esempio di intervento clinico di disostruzione bronchiale condotto con un laser a Nd-YAG. La radiografia in alto a sinistra visualizza una atelettasia del polmone sinistro da ostruzione neoplastica, mostrata nella fotografia endoscopica sotto. L'immagine in basso a destra evidenzia la vaporizzazione del tumore con il laser, con conseguente ricanalizzazione del bronco, confermata dalla riespansione del polmone visibile nella radiografia in alto a destra. Interventi simili vengono effettuati presso l'Istituto di clinica chirurgica generale di Bari e l'Istituto dei tumori di Milano e, in un prossimo futuro, presso la Divisione di chirurgia toracica dell'ex Arcispedale di S. M. Nuova di Firenze.

del tumore colpito; si procede così fino alla completa esposizione del tumore. La necrosi coagulativa che si verifica conduce alla distruzione del tumore, che spesso si stacca spontaneamente dalla parete della vescica. In via alternativa è possibile procedere con l'elettrocoagulazione endoscopica tradizionale e il successivo irraggiamento del letto del tumore per distruggere eventuali cellule neoplastiche e impedire le possibili recidive.

Rispetto alle tecniche convenzionali (in particolare all'elettrocoagulazione), la fotocoagulazione laser presenta, comunque, i vantaggi di un trattamento senza contatto e un migliore controllo di penetrazione e della necrosi tissutale.

Tra le molte applicazioni in dermatologia ci limitiamo a ricordare il trattamento di emangiomi *(port wine stain)* del volto. Queste anomalie vascolari sono costituite da vasi dilatati contenuti prevalentemente nel primo millimetro di pelle, di colore rosso o purpureo. Il trattamento col laser ad argo consente di ottenere, in un'alta percentuale di casi, una riduzione delle dimensioni dei vasi e del loro contenuto di eritrociti. Poiché da ciò dipende, in larga misura, il colore della lesione, ne risulta un miglioramento dell'estetica del volto.

Una menzione particolare merita infine l'impiego del laser nella fototerapia dei tumori con ematoporfirina, una sostanza fotosensibilizzante appartenente all'ampia classe delle porfirine. Questa terapia, che si basa su fenomeni parzialmente osservati fin dai primi decenni di questo secolo, si è sviluppata a partire dal 1975, principalmente per opera di un gruppo di ricercatori del Roswell Park Memorial Institute di Buffalo, nello stato di New York. In essa il laser svolge un ruolo importante non solo come mezzo terapeutico, ma anche come potente strumento di indagine dei fenomeni su cui la terapia si basa. L'ematoporfirina o meglio un suo derivato, che brevemente chiameremo HpD (Hematoporphyrin Derivative), possiede la seguente importante proprietà. Sciolto con opportuno dosaggio in soluzione fisiologica e somministrato per endovena a un paziente affetto da tumore, si localizza dopo 48-72 ore in modo preferenziale nel tumore stesso. Più precisamente, dopo questo tempo, il tumore contiene HpD con una concentrazione di regola 4-5 volte superiore rispetto al tessuto normale circostante. Se il tumore viene illuminato con luce rossa (a 631 nanometri) quando possiede un alto con-

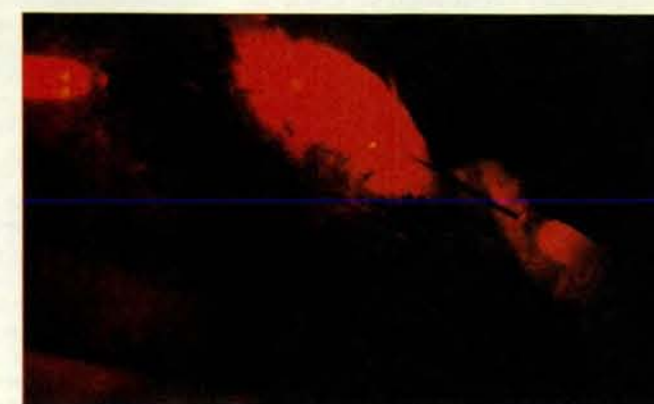
In grafico viene riportato l'andamento temporale della fluorescenza eccitata da un impulso laser da 150 picosecondi in un linfocita umano artificialmente stimolato in fase G_1 *(curva in colore)*, e in uno normale in fase G_0 *(curva in grigio)*, trattati con il derivato dell'ematoporfirina (HpD). La curva che si riferisce ai linfociti stimolati è stata ridotta di un fattore 4. Il confronto tra le due curve è stato fatto per studiare la dipendenza della fluorescenza dallo stato funzionale delle cellule. Le curve sono costituite dalla sovrapposizione di diversi decadimenti esponenziali: 3 per la curva in colore e 2 per quella in grigio. La differenza nei tempi di decadimento delle due curve indica la presenza di diversi costituenti fluorescenti nelle due cellule, oppure una diversa interazione di tali costituenti con l'ambiente cellulare. Il fatto che la curva in colore possieda un'ampiezza di picco 4 volte maggiore indica, inoltre, che la quantità di molecole fluorescenti è diversa nei due casi; con un'eccitazione continua l'intensità di fluorescenza sarebbe proporzionale all'area sottesa da ciascuna curva; le differenze precedenti non sarebbero perciò rilevabili. L'individuazione delle condizioni di massima differenza di emissione tra cellule normali e cellule tumorali consente di migliorare la sensibilità di questa tecnica nella localizzazione, e potenzialmente nella diagnosi precoce dei tumori. Questa ricerca è svolta in collaborazione dal Centro di elettronica quantistica e strumentazione elettronica del CNR di Milano e dal Centro di istochimica del CNR di Pavia, nell'ambito del Progetto finalizzato «Laser di potenza» del CNR.

tenuto di HpD, allora la sua parte illuminata va rapidamente in necrosi e può considerevolmente regredire.

Il processo primario cui è dovuta l'azione combinata dell'HpD e della luce sui tumori sembra ormai accertato: l'HpD che assorbe luce, trasferirebbe poi parte dell'energia assorbita all'ossigeno presente nelle cellule. Tale energia porterebbe l'ossigeno in uno stato molto reattivo (stato eccitato di singoletto) da cui inizierebbe una serie di reazioni che provocano la morte delle cellule. Questi effetti, ormai verificati su molti tipi di tumori, anche se non ancora corredati da una statistica ampia, né da un'osservazione su un periodo di tempo adeguato, né da studi clinici controllati, sono tuttavia sufficientemente interessanti da suggerirne un uso terapeutico. Ma non è tutto; l'HpD è anche fluorescente: illuminata nel violetto (attorno ai 400 nanometri) emette luce rossa. Questa emissione fluorescente, che ovviamente è circoscritta alla regione che contiene l'HpD, cioè al tumore, può essere utilmente impiegata per la localizzazione e la diagnosi precoce.

Perché usare il laser in queste applicazioni? Perché, se il tumore non è superficiale, può essere illuminato con una o più fibre ottiche che raggiungono il tumore tramite endoscopi convenzionali o, molto spesso, tramite semplici aghi da siringa inseriti nel tumore o nelle sue vicinanze.

Ampie e approfondite ricerche sono state compiute, negli ultimi anni, in questo campo, in numerosi istituti specializzati, ma nonostante la gran mole di lavoro svolto diversi aspetti di questa terapia non sono ancora del tutto noti; non si conoscono infatti le ragioni dell'accumulo preferenziale dell'HpD nei tumori; le reazioni che provocano la morte delle cellule; e la stessa composizione dell'HpD, in cui coesistono diverse forme di aggregati con diverse proprietà.

Quanto è stato esposto nell'articolo è frutto dell'esperienza acquisita nell'ambito del Progetto finalizzato «Laser di potenza» del CNR, che ha contribuito alla diffusione delle tecniche laser, operando un'approfondita analisi del quadro di utilità clinica del laser e promuovendo lo sviluppo industriale in tale settore.

Trattamento di un mastocitoma di cane con HpD. L'illuminazione viene fatta inviando la luce di un laser a rodamina B (631 nanometri) in fibre ottiche inserite nel tumore tramite aghi. La scelta di questa lunghezza d'onda è dovuta alla presenza di un picco di assorbimento dell'HpD a 630 nanometri e alla maggiore penetrazione della luce rossa rispetto a quella blu o violetta. Il trattamento è stato eseguito in collaborazione tra la Clinica chirurgica veterinaria dell'Università di Milano e il Centro di elettronica quantistica e strumentazione elettronica del CNR.

L'origine del fondo cosmico a raggi X

In base ad alcune nuove scoperte, una delle sorgenti del fondo di radiazione X che sembra riempire uniformemente l'universo potrebbe essere costituita da un gran numero di quasar lontani

di Bruce Margon

Una delle osservazioni fondamentali sulla struttura dell'universo può essere eseguita a occhio nudo in qualsiasi notte serena: lo spazio tra le stelle è nero. Il fatto che la maggior parte della radiazione visibile nel cielo notturno provenga da singole stelle e che queste sorgenti luminose siano a loro volta raggruppate in galassie è per la cosmologia una delle acquisizioni fondamentali. Un osservatore che esamini il cielo nelle regioni ultravioletta, infrarossa o radio dello spettro elettromagnetico constata tra gli oggetti che costituiscono l'universo la stessa tendenza al raggruppamento. Sulle lunghezze d'onda, molto minori, della radiazione X, però, la situazione appare completamente diversa: fin dal 1962 si sa che ogni zona del cielo emette una «luminosità» uniforme di raggi X, ma dopo venti anni di intense ricerche l'origine di questo fondo di raggi X diffuso è ancora controversa.

Il cielo risulta uniformemente luminoso solo su un'altra banda di lunghezze d'onda: nella regione delle microonde, e per una strana coincidenza della storia dell'astronomia il fondo di radiazione a microonde è stato scoperto nel 1965, subito dopo la prima osservazione del fondo a raggi X. Nel frattempo l'origine del fondo a microonde è stata determinata, e questo fenomeno è pressoché universalmente riconosciuto come un'importantissima chiave per la comprensione dell'evoluzione dell'universo: la maggior parte degli astrofisici considera il fondo a microonde un residuo del big bang, la grande esplosione che segnò l'inizio della fase di espansione dell'universo attualmente in corso.

Le lunghezze d'onda della radiazione X e delle microonde differiscono per un fattore dell'ordine del milione: inoltre questi due tipi di radiazione elettromagnetica vengono normalmente prodotti in natura in condizioni fisiche molto diverse: per questo non c'è motivo di attendersi nessun collegamento diretto tra i due tipi di radiazione cosmica di fondo. Ciononostante l'emissione a microonde è per gli studiosi del fondo a raggi X un suggerimento stimolante, da un lato, e una fonte di continue frustrazioni, dall'altro: lo stimolo consiste nella possibilità che la comprensione dell'origine del fondo a raggi X, pur se ancora lontana, possa un giorno fornire strumenti concettuali per la comprensione della struttura dell'universo su vasta scala, potenti quanto quelli offerti dal fondo a microonde; la frustrazione deriva dal fatto che, nonostante lo studio dei due fenomeni abbia avuto inizio nello stesso periodo, il fondo a microonde è stato ormai spiegato abbastanza esaurientemente, mentre per quanto riguarda il fondo a raggi X sono ancora in discussione gli stessi meccanismi fisici alla base della sua produzione.

La disparità dei progressi ottenuti nello studio dei due tipi di radiazione di fondo si può attribuire in parte alle diverse tecniche di osservazione impiegate alle due lunghezze d'onda: mentre le microonde, la cui lunghezza d'onda va da 0,1 a 10 centimetri, attraversano agevolmente l'atmosfera terrestre, e si possono quindi rivelare con antenne a Terra, nella regione dello spettro dove è stato scoperto il fondo a raggi X, tra 0,1 e 1 nanometro (cioè tra 10^{-10} e 10^{-9} metri), l'atmosfera terrestre risulta completamente opaca. Per le osservazioni astronomiche è quindi necessario porre il rilevatore al di sopra dell'intera atmosfera. Furono i razzi V-2 tedeschi catturati dagli Stati Uniti alla fine della seconda guerra mondiale a offrire agli astronomi americani la prima occasione di portare i loro strumenti nello spazio: un rilevatore di raggi X piuttosto primitivo in uno di questi primi voli mostrò che la corona solare, lo strato più esterno del Sole, estremamente caldo e diluito, è una sorgente identificabile di radiazione X.

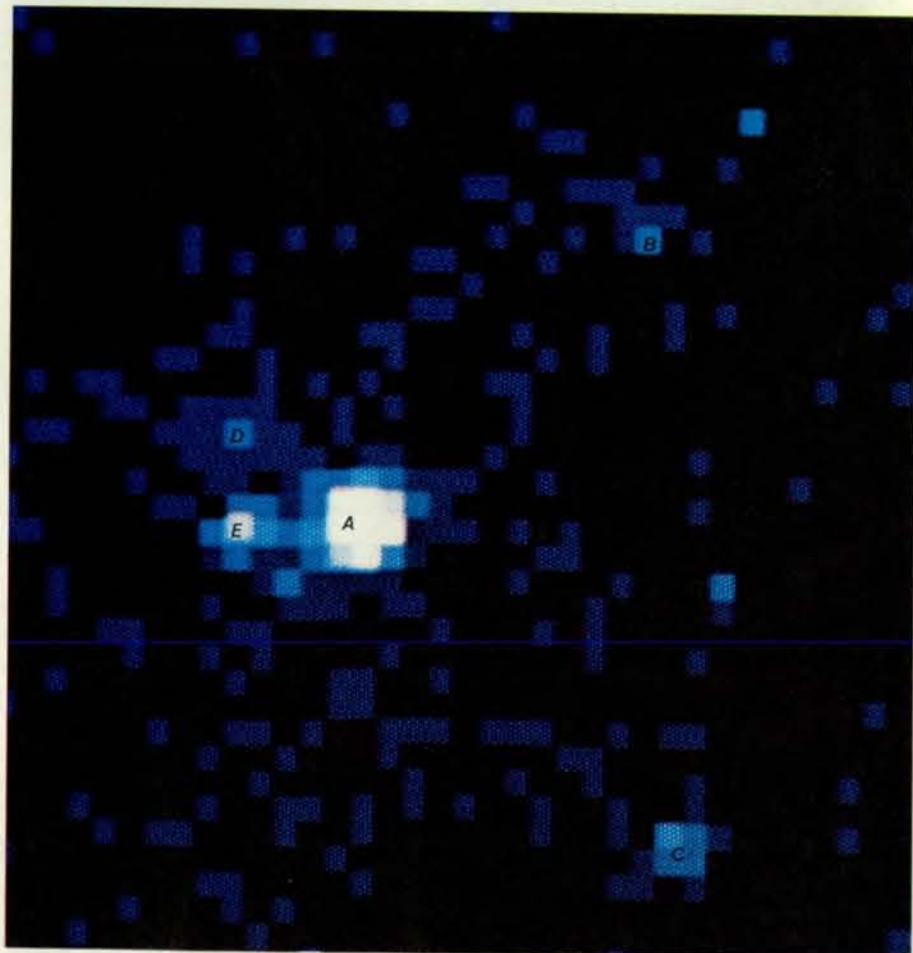
In un certo senso la scoperta di raggi X provenienti dal Sole fu un vero trionfo, poiché segnò la nascita di una nuova tecnica di osservazione astronomica, ma da un altro punto di vista costituì uno scacco: l'emissione solare di raggi X è estremamente debole, circa un milionesimo dell'emissione di luce visibile. Tramite questo rapporto, e sapendo che il Sole è una stella normale, si può valutare facilmente l'emissione che ci si può attendere dai cento miliardi di stelle circa della nostra galassia: i calcoli mostrano che anche la più vicina delle stelle prossime al Sole dovrebbe inviare alla Terra una radiazione 40 miliardi di volte più debole. In altre parole se si rilevano raggi X provenienti dal Sole non è perché questo sia un oggetto notevole, ma solo perché è molto vicino: la prima conquista dell'astronomia X, il rilevamento di raggi X provenienti dal Sole, sembrava quindi suggerire che l'emissione X sarebbe stata di utilità molto scarsa nell'astronomia stellare, almeno per quanto riguarda lo studio della grande maggioranza delle stelle normali.

Fortunatamente Riccardo Giacconi e i suoi colleghi della American Science & Engineering, Inc., non furono dissuasi da questa considerazione dall'indagare il ruolo potenziale dell'osservazione X nello studio degli oggetti celesti anomali: nel 1962 costruirono un rilevatore di raggi X migliorato, che fu sollevato per breve tempo al di sopra dell'atmosfera da un piccolo razzo. Il nuovo rilevatore era decisamente più sensibile di quelli precedentemente impiegati per lo studio del Sole e fu il primo dispositivo nello spazio che avesse una sensibilità sufficiente a rilevare sorgenti di raggi X al di fuori del sistema solare, nel caso fossero esistite.

L'esperimento di Giacconi portò a due scoperte eccezionali: in primo luogo venne individuata una sorgente puntiforme di raggi X, in direzione della costellazione dello Scorpione; la radiazione X che ne proveniva, anche se un milione di volte più debole di quella proveniente dal Sole, era pur sempre molto più intensa di quella prevista per la più vicina delle stelle normali. Questa prima osservazione aveva determinato solo approssimativamente la posizione dell'oggetto appena scoperto, che fu chiamato Sco X-1, e quindi non fu possibile stabilire se la sorgente dell'emissione X corrispondesse a un oggetto visibile. I successivi esperimenti attuati da gruppi di astronomi di vari paesi hanno individuato parecchie centinaia di oggetti che emettono raggi X all'interno della Galassia e li hanno analizzati con sempre maggior precisione, tanto che oggi il meccanismo di funzionamento di queste sorgenti è sufficientemente chiaro.

La previsione che non sarebbe stato possibile rilevare raggi X se non da poche tra le stelle normali più vicine si rivelò esatta. Le sorgenti X come Sco X-1 sono sistemi stellari molto più lontani e decisamente anomali: la radiazione che emettono viene prodotta in un processo che comporta lo scambio di materia stellare tra una stella abbastanza comune e una stella compatta in uno stadio avanzato della propria evoluzione che orbita molto vicino alla prima; nei sistemi binari di questo tipo la compagna compatta può essere una nana bianca, una stella di neutroni e forse anche un buco nero.

I pochi minuti di dati desunti da questo

I quasar e le stelle normali sembrano molto simili nelle comuni fotografie in luce visibile *(in alto)*, ma si possono distinguere facilmente con l'aiuto della corrispondente immagine a raggi X *(in basso)*. La fotografia, riprodotta qui in negativo per una migliore resa dei particolari, è stata ottenuta con il telescopio Schmidt da 120 centimetri di Monte Palomar e rappresenta una regione di cielo di circa mezzo grado di lato (approssimativamente il diametro angolare della Luna piena) nella costellazione di Ercole. La fotografia contiene circa 1000 immagini dall'apparenza stellare, gran parte delle quali sono stelle della nostra galassia; per questo motivo qualsiasi tecnica selettiva progettata per individuare i pochissimi quasar che devono trovarsi in una regione di questo genere deve essere molto efficace nell'eliminazione delle stelle in primo piano. L'immagine X a falsi colori della stessa zona è ricavata dai dati provenienti da un rilevatore di raggi X a bordo del satellite orbitante denominato Osservatorio Einstein; vi si distinguono solo cinque sorgenti di radiazione X, quattro delle quali si possono associare ad altrettanti quasar. L'emittente indicata con *A* è un quasar già noto in precedenza, che era il soggetto originale di questa immagine: è denominato 3C 345 e la sua distanza dalla Terra è stimata in 7,2 miliardi di anni luce. Le altre quattro sorgenti X sono state scoperte proprio grazie a questa immagine; basandosi su osservazioni spettroscopiche compiute con l'aiuto di potenti telescopi, l'autore e i suoi colleghi sono riusciti a stabilire che le sorgenti scoperte fortunosamente, indicate con *B*, *C* e *D*, sono in realtà quasar mai catalogati prima. Le emittenti *B* e *C* si trovano rispettivamente a 9,5 e a 6,7 miliardi di anni luce di distanza, mentre la sorgente *D* è la combinazione delle emissioni di due quasar distinti, uno alla distanza di 7,9 miliardi di anni luce e l'altro a 7,2 miliardi di anni luce. La sorgente *E* coincide con la galassia NGC 6212, relativamente vicina, ma anomala, che si trova a soli 560 milioni di anni luce circa da noi. Nella fotografia sono indicate le controparti ottiche delle varie emittenti X. L'immagine a raggi X è stata fornita da Gary A. Chanan della Columbia University.

primo esperimento «extrasolare» nella regione X dello spettro fornirono un'altra sorpresa, oltre alla scoperta di Sco X-1: il razzo che trasportava il rilevatore per raggi X aveva percorso la sua breve traiettoria ruotando rapidamente su se stesso, spazzando così buona parte della sfera celeste; contrariamente alle previsioni il rilevatore registrò un'emissione di intensità praticamente costante in ogni momento delle osservazioni, indipendentemente dalla direzione in cui queste avvenivano. La scoperta del fondo di radiazione X diffuso aveva quindi luogo contemporaneamente alla nascita dell'astronomia X stellare.

Il fondo a raggi X, come le sorgenti X stellari, è stato oggetto di studio in quasi tutti i successivi esperimenti di astronomia X: mentre i primi tentativi facevano ancora uso di razzi, da circa dieci anni a questa parte i rilevatori vengono montati a bordo di satelliti. Da questi studi sono emerse parecchie caratteristiche affascinanti del fondo di radiazione X; quella forse più notevole è l'uniformità dell'intensità della radiazione proveniente da zone diverse del cielo: nei limiti della migliore accuratezza sperimentale oggi ottenibile, ogni regione mostra una luminosità X esattamente uguale a quella delle regioni circostanti. (Le osservazioni più recenti avrebbero potuto rilevare variazioni di luminosità di circa l'1 per cento tra aree del cielo adiacenti con un'estensione angolare pari all'incirca a quella della Luna piena.) In altri termini il fondo a raggi X è estremamente isotropo: qualunque sia, il meccanismo che lo genera produce la stessa quantità di raggi X in ogni direzione del cielo, con una accuratezza del 99 per cento. (Si sa che anche il fondo di radiazione a microonde è isotropo, e con un'approssimazione ancora maggiore: circa una parte su 3000).

L'isotropia quasi perfetta del fondo a raggi X è evidentemente un indizio fondamentale per comprendere la sua origine; di solito lo si interpreta come una prova che la fonte della radiazione si trova o molto vicino o molto lontano da noi, ma non a una distanza intermedia. È facile capire come sia possibile che la sorgente (o le sorgenti) sia molto vicina: un osservatore completamente immerso nella radiazione emessa da una sorgente a lui prossima registrerebbe una luminosità perfettamente uniforme; nel caso del fondo a raggi X, però, è difficile ipotizzare l'esistenza di una qualsiasi fonte di raggi X abbastanza uniforme nelle vicinanze del sistema solare.

Per capire perché l'alternativa debba proprio essere un'origine della radiazione molto lontana da noi, consideriamo il caso di una ipotetica distribuzione di emittenti X su un intervallo di distanze intermedie, per esempio a distanze paragonabili a quelle che intercorrono tra le stelle nella nostra galassia, o anche tra galassie vicine; in questa situazione dovrebbero esistere in alcune direzioni una o più emittenti più vicine delle altre al sistema solare: guardando in quelle direzioni si dovrebbe vedere una chiazza luminosa corrispondente alla sorgente di radiazione più vicina.

Questa conclusione, però, contraddirebbe l'osservazione iniziale secondo la quale il fondo a raggi X è altamente isotropo: a quanto sembra in nessun punto del cielo si trovano agglomerati di materia che emettono raggi X; l'unica spiegazione possibile è che la maggior parte della radiazione provenga da distanze grandissime, paragonabili a quelle dei più lontani oggetti noti. Come è avvenuto per il fondo a microonde, la spiegazione definitiva del fondo X diffuso avrà quasi certamente qualche conseguenza relativa al problema cosmologico della struttura e dell'evoluzione dell'universo su larga scala.

Il rapporto tra l'energia elettromagnetica rilevata da un ipotetico osservatore extragalattico e quella che si ritiene provenga dalla Galassia è rappresentata tramite barre di colore diverso per le varie parti dello spettro: per sette delle nove bande spettrali considerate il rapporto è minore di 1, il che significa che a queste lunghezze d'onda l'energia presente per unità di volume è più grande all'interno della Galassia che all'esterno. In altre parole l'osservatore esterno valuterebbe che le sorgenti più importanti di questo tipo di radiazione sono concentrate all'interno delle galassie e che l'emissione proveniente dallo spazio intergalattico risulta secondaria. Ma su due gamme di lunghezze d'onda vale l'affermazione inversa: l'energia della radiazione che riempie l'universo è maggiore di quella entro la Galassia; in questo caso l'osservatore esterno troverebbe difficile distinguere le singole galassie, la cui luminosità sarebbe superata da quella del fondo di radiazione diffuso. A differenza della radiazione di fondo a microonde, che viene interpretata come un residuo del big bang, il fondo a raggi X è ancora un mistero. Il diagramma è stato elaborato da Joseph Silk dell'Università della California a Berkeley.

Un secondo indizio sull'origine del fondo a raggi X deriva dal fatto che viene prodotta proprio radiazione X, e non qualche altra radiazione di lunghezza d'onda maggiore: i raggi X sono una delle forme più energetiche di radiazione elettromagnetica; la sorgente dell'emissione osservata deve perciò essere un processo estremamente energetico. Per di più, il processo dev'essere tra quelli che agiscono nelle condizioni tipiche sulla scala dell'intero universo.

I fisici conoscono da decine d'anni, un processo di questo genere, in cui è in gioco quel tipo di gas estremamente caldo e rarefatto che è detto plasma; in questo gas gli elettroni e i protoni, che di solito si trovano legati e formano gli atomi, compaiono come una miscela di particelle libere in rapido movimento: dato che l'elettrone e il protone hanno cariche opposte, tra le coppie di elettroni e di protoni vicini si instaura una forza attrattiva e accade spesso che un elettrone libero, attirato da un protone anch'esso libero, vi passi molto vicino senza però urtarlo. Da quando è nata la fisica quantistica, all'inizio del secolo, si sa che un elettrone libero che transita vicino a un protone libero perde una certa quantità di energia, che si allontana sotto forma di un fotone, un pacchetto discreto di energia elettromagnetica; la radiazione così prodotta si chiama *Bremsstrahlung* o radiazione di frenamento.

L'esatta lunghezza d'onda (o la gamma di lunghezze d'onda) dei fotoni creati dal processo di frenamento dipende strettamente dalla temperatura del plasma. Per un plasma che si trovi a una temperatura compresa tra i 10 e i 500 milioni di kelvin, la radiazione di frenamento dovrebbe consistere di raggi X di lunghezze d'onda per lo più comprese tra 0,1 e 1 nanometro. Esistono motivi per ritenere che l'universo sia pieno più o meno uniformemente di un plasma estremamente caldo

L'estrema isotropia, o uniformità, del fondo a raggi X impone che questa radiazione abbia origine o molto vicino o molto lontano dalla Terra, ma non a distanza intermedia. Infatti, se tutte le sorgenti sono vicine *(a)*, un osservatore si troverebbe completamente immerso nella radiazione e percepirebbe un bagliore uniforme, mentre se si trovano tutte lontane *(b)* le differenze di distanza fra l'osservatore e le singole sorgenti sarebbero relativamente piccole rispetto al valore medio, cosicché non si osserverebbero chiazze più luminose dello sfondo. Se invece la radiazione avesse origine a una serie di distanze intermedie *(c)*, diventerebbe significativa la differenza tra una sorgente vicina e una lontana e ci si potrebbe aspettare di vedere una macchia più luminosa in corrispondenza della sorgente più vicina. Diversi ordini di considerazioni tendono a far preferire l'ipotesi di un'origine lontana del fondo a raggi X diffuso, nonostante sia ancora in dubbio se la sorgente sia costituita o da un gas omogeneo, o da un insieme di oggetti discreti, oppure da una qualche combinazione dei due tipi di emittenti.

in grado di produrre come Bremsstrahlung la radiazione che oggi osserviamo come fondo a raggi X? Vi sono numerose ragioni per sospettare l'esistenza di un gas proprio con queste caratteristiche.

Lo spazio tra le stelle, pur essendo un vuoto estremamente spinto per gli standard terrestri, contiene una quantità non trascurabile di gas molto rarefatto: questo mezzo interstellare ha una massa pari a non più del 10 per cento circa di quella complessiva delle stelle della Galassia, ma svolge un ruolo fondamentale nell'evoluzione galattica costituendo la materia prima da cui si formano nuove stelle. Da molto tempo gli astronomi si sforzano di comprendere per analogia come si formino i giganteschi agglomerati di materia da cui nascono le singole galassie; oggi sembra probabile, anche se tutt'altro che certo, che le galassie si siano condensate da un gas rarefatto che riempiva tutto l'universo poco dopo il big bang: il mezzo intergalattico.

È necessario ammettere la natura ipotetica di questa rappresentazione: mentre il mezzo interstellare può essere osservato direttamente e si sa che la formazione di stelle è un processo ancora in atto anche nella nostra galassia, il mezzo intergalattico non è stato ancora osservato e la formazione delle galassie sembra essersi arrestata da miliardi di anni in tutto l'universo. Eppure c'è motivo di sospettare che un qualche mezzo intergalattico esista ancora oggi: pochi processi naturali hanno un'efficienza del 100 per cento; sarebbe strano, perciò, che la formazione delle galassie abbia assorbito ogni singolo atomo del mezzo primordiale.

Se il fondo a raggi X diffuso va interpretato come una radiazione di frenamento proveniente da un mezzo intergalattico sorge un altro problema: se sia ragionevole presumere che un plasma di questo tipo raggiunga le temperature estremamente elevate necessarie alla produzione di raggi X. Vi sono due ragionamenti che mostrano come questa ipotesi possa essere plausibile, il primo dei quali opera per esclusione. Posto che esista un mezzo intergalattico, se questo gas fosse freddo come quello tra le stelle darebbe luogo, frapponendosi tra la nostra galassia e gli oggetti lontani, a caratteristiche di assorbimento rilevabili nella regione visibile dello spettro di quegli oggetti. Queste caratteristiche di assorbimento, che, come si può dimostrare, hanno tratti molto peculiari, sono state oggetto di attente ricerche, ma risultano in generale assenti: il gas intergalattico freddo, se mai esiste, non è certo sufficientemente ben distribuito da riempire l'universo in modo da apparire isotropo. Da questo fatto, per esclusione, si è indotti a sospettare l'esistenza di un gas intergalattico caldo, nel senso che se un gas di questo tipo esiste in quantità non trascurabile, vi sono buone prove osservative in base alle quali non può essere freddo.

Il secondo ragionamento indiziario a favore di un gas intergalattico caldo si basa sulla valutazione di quanto tempo questo gas possa rimanere caldo: se risultasse che il mezzo intergalattico si raffredda in un tempo breve rispetto alle valutazioni attuali dell'età dell'universo, ci si troverebbe nella scomoda situazione di dover non solo spiegare come il gas abbia ottenuto inizialmente una temperatura così incredibilmente alta, ma anche identificare il meccanismo che ha continuato a riscaldarlo durante tutta la vita dell'universo, tanto da rendere il fondo a raggi X osservabile ancora oggi. In realtà è facile calcolare che in un plasma così rarefatto il tempo di raffreddamento risulterebbe paragonabile all'età dell'universo, se non maggiore.

Ciò avviene perché la radiazione di frenamento è un modo poco efficiente di estrarre energia da un mezzo così poco denso e non esistono altri meccanismi di raffreddamento plausibili. Supponendo quindi che in qualche momento dell'evoluzione dell'universo il gas intergalattico sia stato in qualche modo riscaldato, questo gas dovrebbe essere caldo ancora oggi. Le fonti di calore plausibili, attive durante le prime, violente fasi dell'evoluzione dell'universo, sono più di una: anche se la temperatura del gas è molto alta, si può dimostrare che la quantità di gas presente in ogni piccolo volume di universo è con ogni probabilità estremamente piccola. Pertanto il contenuto di energia complessivo del gas (e quindi l'energia che il meccanismo di riscaldamento primordiale ha dovuto fornire) risulta notevole, ma non eccessivo.

C'è anche un precedente per quanto riguarda l'osservazione di quantità non trascurabili di plasma caldo su scala cosmica, anche se non così diffuso come dovrebbe essere un mezzo effettivamente intergalattico. In tutto l'universo, infatti, si nota una tendenza delle galassie a raggrupparsi in grandi ammassi legati dall'attrazione gravitazionale, i più ricchi dei quali possono contare centinaia o anche alcune migliaia di membri, e una delle prime scoperte dell'astronomia X è stata che gli ammassi ricchi tendono a essere sorgenti rilevabili di raggi X. Approfonditi studi dell'emissione X degli ammassi hanno rivelato che questa non è dovuta alla sovrapposizione di sorgenti distinte appartenenti alle singole galassie, ma appare estesa e diffusa a colmare lo spazio tra di esse e occupa spesso una regione di dimensioni paragonabili a quelle dell'intero ammasso. Attualmente si pensa che questa emissione di raggi X sia radiazione di frenamento proveniente da un plasma caldo (a temperature di norma comprese tra i 20 e i 60 milioni di kelvin) il quale riempie tutto il volume dell'ammasso, coesistendo con le galassie che lo compongono.

Dalla luminosità della radiazione X osservata si può calcolare che la quantità di materia che costituisce il mez-

Un ammasso di galassie ricco è una prova dell'esistenza su scala cosmica di un plasma (un gas estremamente caldo di particelle cariche) che emette raggi X. La fotografia in luce visibile riprodotta in negativo in alto mostra l'ammasso Abell 2151 in un'osservazione compiuta al telescopio Schmidt di Monte Palomar. L'ammasso, che si trova nella costellazione di Ercole, è a circa 670 milioni di anni luce dalla Terra. L'immagine X a falsi colori in basso, sempre fornita da Gary A. Chanan, mostra la stessa regione del cielo come è apparsa all'Osservatorio Einstein: invece delle singole galassie che compaiono nell'immagine ottica, quella a raggi X rivela una vasta nube di gas caldo, di circa tre milioni di anni luce di diametro, trattenuta dalla forza gravitazionale dell'agglomerato di galassie. Il mezzo che riempie l'ammasso ha un diametro angolare di 0,25 gradi, ossia pari a circa metà di quello della Luna piena. Un plasma di questo tipo, la cui temperatura si può stimare, grazie alla sua emissione X, nell'ordine dei 20 milioni di kelvin, è stato individuato in molti altri ammassi galattici analoghi.

zo entro l'ammasso, risulta paragonabile alla massa complessiva delle galassie che ne fanno parte; ne deriva che, almeno in questi ammassi ricchi, il plasma caldo non è certo uno stato poco comune della materia, ma anzi una delle sue forme prevalenti. Negli ultimi anni l'osservazione del gas interno agli ammassi ha fornito buoni motivi per ritenere che esso abbia avuto origine all'interno delle galassie; il gas mostra tracce di elementi, come il ferro, che si pensa vengano prodotti solo all'interno delle stelle, ed è quindi probabile che la sua comparsa sia successiva a quella delle galassie e che non sia quindi veramente rappresentativo del mezzo intergalattico caldo. Ciononostante il mezzo interno agli ammassi rimane un esempio concreto del fatto che l'universo, almeno in certe zone, contiene effettivamente grandi quantità di plasma estremamente caldo che emette raggi X.

Accettando per il momento l'ipotesi che il fondo a raggi X diffuso si possa attribuire alla radiazione di frenamento proveniente da un mezzo intergalattico caldo, è possibile che questo plasma, oltre a risolvere un problema osservativo aperto ormai da vent'anni, abbia altre proprietà notevoli? La risposta è decisamente affermativa. Nei primi anni venti Edwin P. Hubble e i suoi colleghi del Mount Wilson Observatory, studiando gli spettri di varie galassie, stabilirono che tutte le galassie al di fuori dell'ammasso locale se ne stanno allontanando: in breve, l'intero universo è in espansione. Questa scoperta portò all'idea che la fase attuale dell'evoluzione dell'universo sia stata preceduta da almeno un'altra fase, violenta e caratterizzata da un'elevata densità; quest'idea venne successivamente confermata dalla scoperta della radiazione di fondo a microonde.

L'attuale fase di espansione durerà per sempre? Se fosse così, tra 10 o 20 miliardi di anni l'universo diventerebbe decisamente monotono: non solo perché tutte le galassie si saranno completamente separate le une dalle altre, ma anche perché l'evoluzione stellare è un processo unidirezionale e limitato che finisce per trasformare tutta la materia disponibile nelle galassie in stelle oscure e compatte, nane bianche, stelle di neutroni, e forse anche buchi neri, così che le galassie finiranno per esaurire le sorgenti della loro luminosità affievolendosi fino a scomparire. Questa triste previsione si basa però sull'ipotesi fondamentale che l'espansione proceda incontrastata; l'esistenza del mezzo intergalattico potrebbe modificarne il destino.

Esiste infatti una forza che agisce contro l'espansione dell'universo, e precisamente l'attrazione gravitazionale tra ciascun atomo del cosmo e tutti gli altri; se l'autogravitazione dell'universo è sufficientemente forte, l'espansione finirà per fermarsi trasformandosi in una contrazione governata dalla gravità. Come è possibile risolvere questo problema cruciale? Dal punto di vista osservativo gli astronomi di oggi sanno solo quello che Hubble scoprì cinquant'anni fa: l'universo si trova attualmente in una fase di espansione. È interessante notare che, per poter prevedere se l'espansione continuerà indefinitamente o se subentrerà una contrazione, non occorrono procedimenti complicati: dato che l'intensità della forza gravitazionale dipende solo dalla quantità di materia in gioco e dalla distanza che separa i vari aggregati di materia, mentre il tasso di espansione e la scala dell'universo in questo momento si possono almeno in linea di principio determinare con osservazioni sull'espansione delle galassie, il problema del futuro dell'espansione cosmica è riducibile al calcolo della quantità complessiva di materia presente nell'universo.

Più esattamente, è la densità dell'universo (cioè la quantità di materia presente in ogni piccolo volume dato) che deve essere calcolata: se questa è superiore a un valore limite, noto come densità critica, l'espansione dovrà aver fine, e seguirà una contrazione; altrimenti, se la densità media dell'universo risulta inferiore a quella critica, l'espansione continuerà per sempre. Attualmente la stima migliore della densità critica fornisce un valore estremamente piccolo, anche perché il volume dell'universo in questo momento è molto grande; una media di un atomo di idrogeno per metro cubo.

Prescindendo dall'esistenza di un mezzo intergalattico, si ritiene che la massima parte della materia dell'universo si trovi nelle galassie, sotto forma di stelle e, in misura minore, di gas e di polvere. Il calcolo della massa delle galassie richiede un procedimento complesso e tortuoso, ma il risultato è talmente significativo, nel contesto della nostra discussione, da rendere abbastanza poco importante le possibili imprecisioni. Si arriva alla conclusione che la massa della galassie porta a una densità pari solo all'1 o al 2 per cento di quella critica; per questo, anche considerando un certo margine di incertezza nella stima della materia galattica, sembra difficile che le galassie normali dispongano di massa sufficiente ad arrestare l'espansione.

L'esistenza del fondo a raggi X diffuso e la sua possibile origine da un plasma intergalattico caldo, tuttavia, impongono un riesame di questi calcoli: se la densità del mezzo intergalattico superasse di un fattore 50 o 100 la densità media dell'universo attribuibile alle galassie, l'universo raggiungerebbe la sua densità critica e le previsioni relative al suo futuro cambierebbero drammaticamente. Per di più ci si verrebbe a trovare in una strana situazione: gli astronomi avrebbero sempre ignorato, fino a oggi, qualcosa come il 99 per cento della massa dell'universo e le belle galassie luminose con tutto il loro contenuto sarebbero retrocesse al rango di impurità presenti in tracce nel mezzo intergalattico!

Se il fondo a raggi X è dovuto al mezzo intergalattico, è facile valutare la densità del mezzo stesso misurando l'intensità della radiazione di fondo osservata (cioè quanta radiazione X raggiunge la Terra in ogni secondo da una data direzione) e calcolando teoricamente quanti atomi in ogni volume di spazio sono necessari per creare questa radiazione come radiazione di frenamento. Il risultato di questo calcolo si rivela per il ricercatore una vera tortura. Se il fondo a raggi X è dovuto alla radiazione di frenamento di un plasma caldo, allora, a seconda del valore che si attribuisce a parecchie grandezze poco note, la densità ricavata per il plasma intergalattico può benissimo essere vicina alla densità critica dell'universo.

La scoperta del fondo a raggi X ha modificato le nostre idee sulla futura espansione dell'universo? Purtroppo, se la radiazione di fondo è un indizio di una densità critica dell'universo, non ne è però una prova. La catena del ragionamento è interrotta in almeno tre punti fondamentali. Innanzitutto il valore della densità critica dipende strettamente dalla velocità di espansione dell'universo: più veloce è la recessione delle galassie le une dalle altre, più materia occorre per limitare l'espansione. Le stime di questa velocità, rappresentata dalla grandezza nota come costante di Hubble, derivano da tutta una serie di complesse misurazioni astronomiche dominate da diverse incertezze. In realtà dai tempi di Hubble la valutazione accettata di questa costante ha continuato a cambiare in maniera notevole, mentre il valore ricavato della quantità di mezzo intergalattico sarebbe la densità critica della materia solo per una ben determinata grandezza di questa costante, per di più abbastanza lontana da quella oggi ritenuta corretta.

Un secondo motivo di incertezza è che la determinazione precedentemente discussa della densità del plasma, essendo fondata sull'osservazione dell'intensità dei raggi X, richiede la conoscenza della granulosità media della regione che li emette, poiché un mezzo uniforme e uno granuloso emettono radiazioni con efficienza leggermente diversa. Solo se il mezzo intergalattico è estremamente uniforme la densità dedotta è pari a quella critica e anche se non è possibile misurare direttamente il grado di granulosità, dal punto di vista teorico un mezzo sufficientemente uniforme sembra piuttosto improbabile: quasi tutti i costituenti dell'universo, come la materia luminosa contenuta nelle galassie, sono distribuiti in modo decisamente irregolare.

Il terzo dubbio, sulla conclusione che l'universo sia gravitazionalmente legato da un mezzo intergalattico caldo, è forse il più profondo: non esistono prove dirette che il fondo diffuso a raggi X che si osserva sia effettivamente dovuto a questo plasma caldo e non a un meccanismo completamente diverso. Purtroppo l'emissione X non possiede alcun segno distintivo inequivocabile della sua origine come radiazione di frenamento proveniente da un gas rarefatto e d'altra parte esistono altri procedimenti di laboratorio ben noti in grado di generare raggi X; la spiegazione del fondo a raggi X mediante un plasma caldo è compatibile con le osservazioni compiute finora, e forse è anche elegante, nella sua semplicità, ma è davvero l'unica possibile? Data l'importanza cosmologica dell'esistenza del plasma, almeno se la sua densità è vicina a quella critica, si dovrebbe esigere dalla spiegazione qualcosa di più della semplice compatibilità con i dati sperimentali. Sarebbe preferibile, per esempio, sapere con certezza che non esistono per il momento altre spiegazioni del fondo a raggi X compatibili con i dati.

Purtroppo non è così. Esiste una spiegazione completamente diversa del fondo a raggi X diffuso, assolutamente indipendente dall'esistenza di un mezzo intergalattico, e questa spiegazione concorrente della radiazione di fondo è oggetto da alcuni anni di un interesse sempre crescente da parte di coloro che lavorano in questo campo. L'osservazione che il fondo a raggi X è estremamente isotropo dà una misura non solo di una proprietà della sorgente di questa radiazione, ma anche del livello di raffinatezza della strumentazione disponibile. Consideriamo per esempio il caso di un gran numero di oggetti discreti che emettano raggi X: se questi oggetti fossero abbastanza lontani dall'osservatore e tanto numerosi che ogni zona del cielo ne contenesse molti, l'osservatore potrebbe interpretare erroneamente questo insieme di sorgenti discrete come una luminosità indistinta e uniforme; per quanto uno strumento possa essere accurato sarebbe sempre ipotizzabile una qualche densità delle sorgenti nel cielo tale che lo strumento non sia in grado di risolverle come oggetti discreti e di determinare così che non si tratta di un mezzo veramente diffuso, ma di un insieme di punti.

Il fondo a raggi X diffuso proviene davvero da una sorgente diffusa o deriva piuttosto da un gran numero di emittenti X lontane da noi, ma distinte, del tutto indipendenti da un gas intergalattico rarefatto? Un primo passo verso la soluzione di questo problema è chiaramente quello di chiedersi se esista tra i costi-

tuenti dell'universo a noi noti un oggetto molto comune che sia anche una potente fonte di raggi X; le galassie, gli elementi fondamentali della struttura dell'universo, non soddisfano questo duplice requisito. Pur essendo decisamente numerose non emettono raggi X a sufficienza per arrivare all'intensità osservata del fondo di radiazione.

Questa conclusione è a sua volta una conseguenza di quella cui si era giunti subito dopo la prima osservazione di raggi X provenienti dal Sole: le stelle normali non emettono una frazione significativa della loro energia come raggi X e pertanto la somma delle emissioni X di tutte le stelle di una galassia risulta trascurabile in confronto alla sua produzione di luce visibile. Ho già citato un secondo costituente dell'universo che è una notevole sorgente di raggi X: il gas negli ammassi galattici ricchi. Questa sorgente sembrerebbe a prima vista promettente, poiché gli ammassi di galassie sono oggetti relativamente comuni, ma alcune osservazioni recenti dell'emissione X proveniente da ammassi vicini, combinate con una stima della loro densità nel resto dell'universo, suggeriscono che gli ammassi possano fornire tutt'al più una piccola percentuale del fondo a raggi X diffuso.

Negli ultimi tre anni le osservazioni in X compiute dal satellite noto come Osservatorio Einstein hanno rivelato un'altra classe di oggetti che potrebbero essere le sorgenti del fondo di radiazione X: secondo questi dati gli oggetti quasi stellari, o quasar, sono forti emittenti di raggi X. La vera natura dei quasar rimane ancora oggi un enigma. Sulle fotografie ottiche appaiono come punti, identici alle normali stelle della Galassia, ma una ventina di anni fa si scoprì che, al contrario delle stelle normali, molti di questi oggetti apparentemente banali sono sorgenti di un'intensa emissione radio.

Questa osservazione stimolò uno studio più approfondito dell'emissione ottica dei quasar, che a sua volta rivelò ben presto che questi oggetti si allontanano dalla Terra a velocità enorme, in alcuni casi superiore al 90 per cento della velocità luce; se si spiegano queste velocità di recessione con l'espansione complessiva dell'universo, che è anche la causa dell'allontanamento delle galassie lontane (un'ipotesi accettata dalla maggior parte degli specialisti, ma non da tutti), si è costretti a concludere che i quasar non sono stelle della nostra galassia, ma i più lontani oggetti visibili di tutto l'universo. Nei casi più estremi si tratta di oggetti situati ben oltre le più lontane galassie; la luce che vediamo oggi è perciò partita poco dopo il big bang. Nonostante la loro distanza prodigiosa, molti quasar sono abbastanza luminosi da essere visibili con un piccolo telescopio da dilettante, e, quindi, la loro luminosità assoluta deve essere enorme, molto superiore a quella delle galassie normali, e il meccanismo con cui producono energia è con ogni probabilità ben diverso dalla fusione nucleare che alimenta le stelle normali.

L'esame di campioni accuratamente selezionati di oggetti deboli registrati su lastre fotografiche ha rivelato che, pur essendo piuttosto rari nelle vicinanze della nostra galassia, i quasar risultano un costituente abbastanza comune delle regioni molto lontane dell'universo: sembra che un effetto di tipo evolutivo abbia causato la scomparsa pressoché totale del fenomeno dei quasar negli ultimi miliardi di anni.

Già prima del lancio dell'Osservatorio Einstein si sapeva, grazie a osservazioni compiute mediante razzi o dai satelliti precedenti, che alcuni quasar sono sorgenti straordinariamente potenti di radiazione X: nei casi più eccezionali un singolo quasar produce nella sola regione X dello spettro 1000 volte più energia di tutti i 100 miliardi di stelle della nostra galassia nel visibile! Eppure, per la grande distanza a cui si trovano quasi tutti i quasar, anche questa incredibile produzione di energia risulta così affievolita, al momento di raggiungere la Terra, che la maggioranza dei quasar non erano accessibili all'osservazione X prima del lancio dell'Osservatorio Einstein.

Il satellite Einstein ha rilevato raggi X con una sensibilità circa 1000 volte superiore a quella degli esperimenti precedenti. Per di più, al contrario dei precedenti rilevatori, simili a contatori Geiger e capaci di registrare quasi solo l'intensità della radiazione X, il rilevatore del satellite Einstein incorporava un nuovissimo sistema di messa a fuoco in grado di formare vere e proprie immagini X, che venivano poi trasmesse via radio a terra in forma digitale. Nei due anni della sua vita attiva il satellite ha registrato immagini X di un gran numero di oggetti interessanti, tra cui parecchie centinaia di quasar; moltissimi dei quasar così osservati sono risultati sorgenti potentissime di radiazione X.

A partire da queste osservazioni nello spettro X relative a un campione limitato di quasar, combinate con i dati sulla densità dei quasar nelle varie zone del cielo ricavati dai più esaurienti studi ottici, è in linea di principio possibile valutare il contributo dei quasar al fondo a raggi X diffuso. Questo calcolo porta alla conclusione affascinante che gran parte, se non tutto il fondo di radiazione X, è effettivamente dovuta al contributo dei quasar lontani individualmente invisibili. Tuttavia, come per ogni ricerca che si basi sull'esame di una piccola frazione di una popolazione per estrapolare le caratteristiche di un gruppo molto più grande, su questo risultato gravano numerose incertezze.

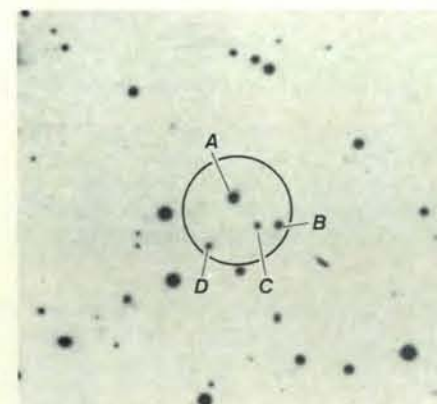
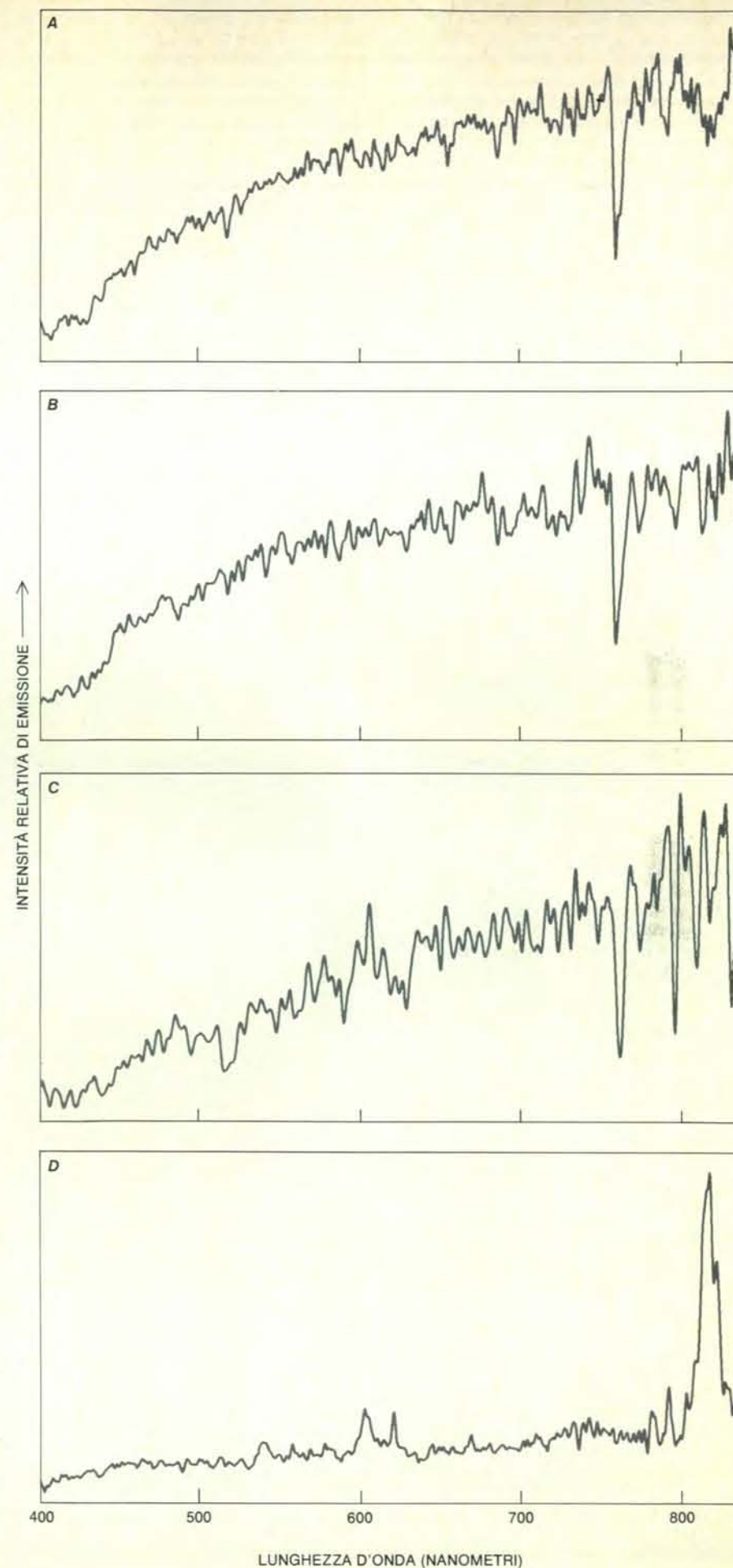
Nel caso dello studio sui quasar che emettono raggi X compiuto dall'Osservatorio Einstein le difficoltà si sommano: innanzitutto gli stessi dati ottici su cui ci si basa per dedurre il numero di quasar per unità di superficie celeste sono ancora parzialmente in discussione, e questi dati sono fondamentali per passare dall'emissione X del piccolo campione di oggetti osservato alla valutazione del contributo complessivo dell'intera categoria. In secondo luogo, come gli astronomi hanno compreso fin dalla scoperta dei quasar, sembra che le caratteristiche di questi strani oggetti si evolvano nel corso della vita dell'universo. In effetti i dati delle osservazioni X permettono perfino di avanzare qualche ipotesi sull'evoluzione delle caratteristiche dell'emissione X dei quasar: secondo alcuni ricercatori del Center for Astrophysics dello Harvard College Observatory e dello Smithsonian Astrophysical Observatory il rapporto tra l'emissione di raggi X e l'emissione nel visibile dei quasar da loro esaminati sembra dipendere da un certo numero di altre proprietà, tra cui forse anche la distanza dei singoli oggetti dalla nostra galassia.

Gli indizi di una simile evoluzione delle proprietà dell'emissione X sono però estremamente difficili da interpretare, soprattutto perché qualsiasi contributo dei quasar al fondo a raggi X deriva necessariamente dalla sovrapposizione delle emissioni provenienti da tutto l'universo e quindi corrispondenti a tutti i diversi stadi della loro evoluzione; un calcolo corretto del contributo dei quasar al fondo X presuppone quindi un'esatta comprensione della natura del processo evolutivo. (Il problema è analogo a quello del calcolo dell'interesse composto su un conto corrente: noto il capitale iniziale e il tasso di interesse è facile prevedere il saldo in ogni momento dato, ma se si utilizza un interesse anche leggermente diverso il saldo previsto varierà notevolmente.)

In ultima analisi il problema dei dati attualmente disponibili è di tipo ben noto a chi si occupa di statistiche e censimenti: fino a che punto è rappresentativo il piccolo campione della popolazione su cui è stato fatto il sondaggio? Tramite l'Osservatorio Einstein sono stati osservati i quasar scoperti negli anni immediatamente precedenti il lancio, perché sono estremamente luminosi o nello spettro visibile o nelle onde radio. Questi oggetti costituiscono la grande maggioranza di tutti i quasar già conosciuti; la domanda è: si tratta di un campione tale da permettere una valutazione corretta delle caratteristiche di emissione X dei quasar?

All'inizio del programma di osservazione con il satellite Einstein l'unica possibilità era quella di studiare i quasar elencati nei cataloghi già esistenti, ma la notevole sensibilità dei suoi rilevatori di raggi X ha finito per offrire un'alternativa imprevista. In un'osservazione tipica, infatti, si orienta il satellite in modo che l'oggetto astronomico in esame si trovasse nel centro dell'immagine X prima dell'inizio dell'esposizione; come per qualsiasi macchina fotografica, nell'inquadratura entra anche un'area circostante, che non ha nulla a che fare con l'oggetto al centro: ebbene, un attento studio delle immagini X ripreso dal telescopio Einstein rivela che in gran parte di queste registrazioni compaiono una o più deboli sorgenti X ulteriori, indipendenti dal soggetto originale della fotografia. Queste nuove emittenti scoperte per un vero e proprio colpo di fortuna del tutto imprevisto sono state battezzate sorgenti Einstein «fortunose» (*serendipitous*).

Più si prolunga il tempo di esposizione di un'immagine a raggi X, più sorgenti fortunose tendono ad apparire su di essa; questa scoperta costituisce un forte indizio del fatto che, se effettivamente qualche tipo di sorgenti puntiformi contribuisce in maniera significativa al fondo X diffuso, le sorgenti fortunose ne costituiscono probabilmente un campione rappresentativo. In altri termini queste sorgenti, comparendo in modo imprevedibile nelle registrazioni, senza alcun intervento da parte dell'osservatore e a sua insaputa, sono la realizzazione del sogno di ogni elaboratore di statistiche: il campione non inquinato per antonomasia.

Qual è la vera natura delle sorgenti fortunose? E soprattutto, esiste una corrispondenza tra questi oggetti e altri già ben

Queste immagini illustrano la scoperta fortunosa di un quasar che emette raggi X. Durante uno studio dell'emissione X proveniente da un sistema stellare binario nella costellazione dell'Acquario, si è scoperta su un'immagine dell'osservatorio Einstein un'altra sorgente X nelle vicinanze di quella originale; la sua posizione risultava determinata solo approssimativamente dai dati dell'osservazione X. Su una fotografia in luce visibile della stessa zona del cielo *(qui sopra)*, ottenuta con il telescopio Schmidt di Monte Palomar, è tracciato un cerchio attorno alla regione in cui sembrava trovarsi la sorgente X; in questo cerchio risultano visibili quattro oggetti dall'aspetto stellare, che sono stati tutti trattati come possibili controparti ottiche della sorgente. Gli spettri di questi oggetti estremamente deboli *(curve a destra)* sono stati ricavati dall'autore e dai suoi colleghi con il riflettore Shane da tre metri del Lick Observatory: i primi tre spettri mostrano profondi avvallamenti alle lunghezze d'onda in cui gli atomi freddi degli strati più esterni dell'oggetto devono aver assorbito la radiazione proveniente dall'interno; queste caratteristiche sono tipiche dello spettro di stelle normali e indicano che questi tre oggetti non sono probabilmente collegati con la sorgente X. Il quarto spettro, invece, ha un aspetto completamente diverso, con picchi pronunciati che corrispondono a emissioni superiori alla media a certe lunghezze d'onda: questo è il marchio caratteristico di un quasar, che in questo caso particolare si trova a circa 3,8 miliardi di anni luce da noi, come si può ricavare dalle caratteristiche dello spettro. Il quasar, denominato 2216-043 dalle sue coordinate celesti, è la controparte ottica della sorgente X fortunosa.

studiati, per esempio con metodi ottici? È difficile rispondere a questa domanda, perché la limitata efficacia dei dispositivi per la messa a fuoco dei raggi X fornisce immagini piuttosto confuse in confronto a quelle ottiche; proprio come è difficile misurare con precisione le caratteristiche di un'immagine sfocata in una fotografia in luce visibile, così è problematico attribuire a ciascuna delle sorgenti fortunose una precisa posizione nel cielo. Tutto quello che si può fare è determinare una piccola regione, dell'estensione di circa un trentesimo di grado, da cui si può ritenere provengano i raggi X, ma questa approssimazione non è sufficiente per identificare univocamente l'oggetto con una sorgente visibile nota: di norma una fotografia in luce visibile della regione del cielo individuata riporta da quattro a sei oggetti debolmente luminosi, ciascuno dei quali (o magari nessuno) potrebbe essere l'immagine ottica della sorgente X.

Per rivelare la natura delle sorgenti X fortunose e per decidere se hanno qualche importanza per quanto riguarda il problema del fondo a raggi X diffuso, sono evidentemente necessarie ulteriori osservazioni, che sono state condotte da vari gruppi di studiosi in tutto il mondo, tra cui uno formato da Gary A. Chanan della Columbia University, Ronald A. Downes dell'Università della California a Los Angeles e da me. La nostra tecnica consiste nell'esaminare delle comuni fotografie in luce visibile della piccola regione celeste che corrisponde all'area di incertezza sulla posizione di ciascuna sorgente fortunosa; come ho già osservato questo procedimento individua di solito pochi oggetti debolmente luminosi ciascuno dei quali è una potenziale controparte ottica della sorgente X. Successivamente esaminiamo questi oggetti visibili alla ricerca di qualche caratteristica anomala (bisogna infatti ricordare che difficilmente le stelle normali sono emittenti rilevabili di raggi X); dato però che questi oggetti sono di norma 100 000 volte più deboli delle stelle visibili a occhio nudo, questo lavoro richiede l'impiego di telescopi particolarmente potenti: noi abbiamo utilizzato soprattutto il riflettore da 2,1 metri del Kitt Peak National Observatory, il telescopio Shane da tre metri del Lick Observatory e quello da quattro metri al Cerro Tololo Inter-American Observatory in Cile.

La nostra tecnica è fondata sulla spettroscopia, la suddivisione di una radiazione nelle lunghezze d'onda che la compongono. Gli spettrogrammi delle stelle normali si distinguono per la presenza di lacune in corrispondenza di certe lunghezze d'onda caratteristiche, dove la luce della stella è stata assorbita, prima di raggiungere il rilevatore, dagli atomi relativamente freddi che si trovano negli strati più esterni della stella stessa. Di norma quasi tutti gli spettri di oggetti stellari vicini alla posizione di una sorgente X fortunosa hanno rivelato una struttura di questo tipo; questi oggetti sembrano perciò stelle normali relativamente vicine a noi che non hanno nulla a che fare con la sorgente X.

Nella maggior parte dei casi, però, siamo riusciti a individuare nella zona considerata un altro oggetto con uno spettro completamente diverso, che presenta un eccesso di emissione ad alcune lunghezze d'onda determinate; una struttura spettrale di questo tipo, esattamente opposta a quella di assorbimento tipica delle stelle normali, indica che si è di fronte a un oggetto decisamente straordinario che dispone evidentemente di notevoli quantità di gas caldo. Per di più le lunghezze d'onda in cui avviene l'emissione corrispondono a quelle di atomi ben noti come l'idrogeno, l'elio e il carbonio, sottoposte però a un sistematico spostamento verso l'estremità rossa dello spettro; questo spostamento verso il rosso mostra che la sorgente si allontana dall'osservatore. Queste proprietà spettrali (intense righe di emissione di elementi chimici comuni, ma spostate verso il rosso) sono il segno distintivo dei quasar. Per farla breve, quello che noi, insieme con vari altri gruppi che lavorano su questo problema, abbiamo scoperto, è che gran parte delle sorgenti di raggi X fortunose è costituita da quasar precedentemente ignoti.

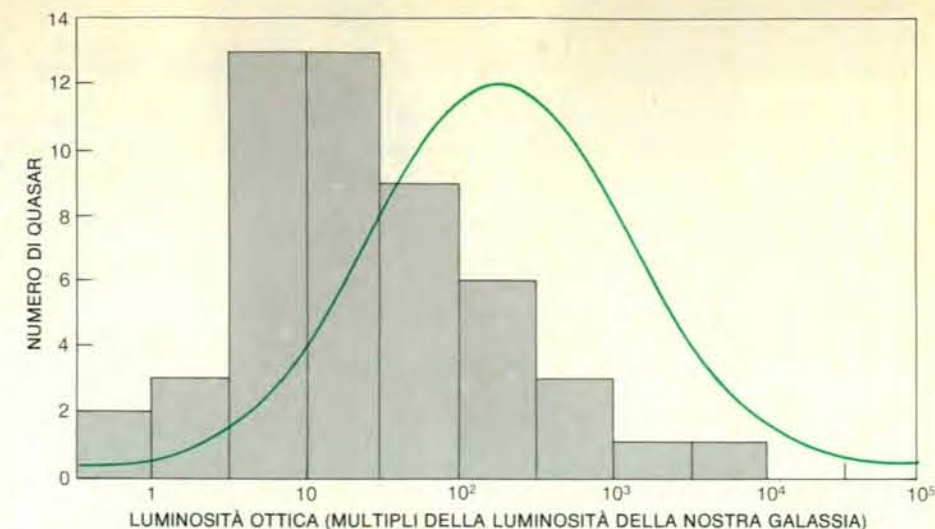
Questa scoperta ha evidentemente ripercussioni profonde sul problema dell'origine del fondo a raggi X diffuso, e in particolare su quelle interpretazioni che attribuiscono la radiazione di fondo alla somma di numerose sorgenti puntiformi lontane. Se, aguzzando gli occhi nella nebbia, si riescono a malapena a identificare gli oggetti più luminosi, all'interno di un gruppo di oggetti lontani che si confondono in un'unica chiazza, sembra ragionevole supporre che gli oggetti di questo tipo siano una parte non trascurabile anche della popolazione più lontana e difficile da studiare. Analogamente, l'identificazione come quasar di un gran numero di sorgenti fortunose (i casi attualmente noti sono circa 150) tende decisamente a suggerire che i quasar siano responsabili di una significativa frazione della radiazione di fondo.

Ma quanto è ampia questa frazione? I quasar sono la fonte più importante del fondo di radiazione, e quindi rappresentano la soluzione dell'annoso problema della sua origine, o costituiscono soltanto una curiosità, un'impurità in un'emissione che invece è dovuta a un plasma intergalattico caldo? Dato che non esiste una stima quantitativa del mezzo intergalattico caldo indipendente dall'intensità del fondo a raggi X, è effettivamente possibile che tutte e due le sorgenti contribuiscano all'emissione X osservata.

La grande maggioranza degli oggetti astronomici emette alle lunghezze d'onda dei raggi X solo una piccola parte della radiazione che produce. Se si attribuisce il fondo a raggi X a qualche tipo di sorgenti distinte lontane, questi oggetti devono essere innanzitutto potenti emittenti X e in secondo luogo devono essere molto comuni per poter spiegare l'isotropia che si osserva. Come indica questo istogramma, è difficile soddisfare contemporaneamente questi due requisiti: le emittenti X più intense tendono anche a essere eccezionali. Le stelle con un rapporto tra emissione X ed emissione visibile maggiore di 1 sono così poche che l'emissione X di una galassia normale è di solito solo un decimillesimo (10^{-4}) della sua emissione nel visibile; per questo le galassie normali, benché numerose, non contribuiscono in maniera significativa al fondo a raggi X diffuso. Il plasma caldo intrappolato negli ammassi di galassie ricchi produce una grande quantità di raggi X, ma gli ammassi di questo tipo sono troppo rari per fornire più di una piccola percentuale della radiazione di fondo. I quasar sono gli oggetti che più si avvicinano al duplice requisito: la loro emissione X ha dell'incredibile e sono numerosi a grandi distanze dalla Terra.

La chiave per risolvere il problema è nelle proprietà di questi quasar scoperti così fortunosamente: uno studio esauriente delle caratteristiche degli elementi di questo campione (i primi quasar individuati per la loro emissione X invece che per quella visibile o radio) dovrebbe fornire un'immagine nuova, e forse meno parziale, delle proprietà dell'emissione X di tutta l'enigmatica classe dei quasar. In effetti un confronto tra le proprietà del campione selezionato per le sue caratteristiche X e quelle di gruppi di quasar precedentemente identificati grazie alla loro emissione visibile o radio rivela alcune differenze interessanti: il valore intrinseco dell'emissione di energia luminosa dei quasar scoperti nello spettro X sembra sistematicamente inferiore, e di un fattore 10, a quella dei quasar già noti.

L'interpretazione precisa di questa differenza, tuttavia, è per ora dubbia: alcuni studiosi ritengono che tanto la luminosità nel visibile quanto quella in X siano ampiamente variabili nell'intera classe dei quasar, cosicché i sondaggi compiuti con metodi ottici individuerebbero le più notevoli emittenti nel visibile, dotate invece di caratteristiche relativamente meno straordinarie nella banda X, mentre i sondaggi in X darebbero risultati opposti. Se così fosse, un confronto tra i due campioni potrebbe benissimo rivelare quelle differenze non trascurabili che effettivamente si osservano. Altrimenti potrebbe accadere che alcuni degli effetti evolutivi citati in precedenza tendano a ridurre l'emissione X soprattutto di un certo gruppo ristretto di quasar, creando ancora una volta una differenza tra le caratteristiche di un campione individuato otticamente e quelle di uno selezionato mediante la sua emissione X. Peggio, ambedue gli effetti potrebbero svolgere un ruolo importante simultaneamente.

La possibilità che agiscano effetti evolutivi è particolarmente preoccupante; si può infatti far vedere che, anche se qualsiasi radiazione isotropa è dovuta alla somma dei contributi provenienti dalle varie distanze, prevalgono quelli degli oggetti più lontani. Ma se le caratteristiche degli oggetti si evolvono e quindi variano con la distanza, quelli più lontani e meno luminosi, sui quali i dati sono scarsi, possono essere molto diversi da quelli più vicini e più luminosi. I dati su questi ultimi, anche se soddisfacenti, non risultano rappresentativi dell'intera popolazione; quasi sicuramente questo effetto diventa importante a qualche livello anche per i quasar che emettono raggi X. Dato poi che persino la sensibilità straordinaria del rilevatore del satellite Einstein non era sufficiente per lo studio degli oggetti che devono essere responsabili della parte principale del fondo a raggi X, se davvero è dovuto principalmente ai quasar, per determinare il contributo complessivo dei quasar alla radiazione diffusa è necessario compiere una estrapolazione delle proprietà osservate dei quasar scoperti nello spettro X. I limiti di validità di queste estrapolazioni sono per ora sconosciuti; in realtà, a seconda del procedimento adottato, si può dedurre che i quasar sono una componente fondamentale del fondo di radiazione oppure che la loro importanza è trascurabile.

Su questo istogramma sono rappresentati 51 quasar, individuati dall'autore e dai suoi colleghi come controparti ottiche di altrettante sorgenti X fortunose, comparse sulle immagini registrate dall'Osservatorio Einstein, in funzione della loro emissione di luce visibile. Prima di questo studio, praticamente tutti i quasar noti erano stati scoperti per le loro insolite proprietà ottiche o radio, e non per la loro emissione X. La curva in colore rappresenta la distribuzione in funzione della luminosità ottica di un campione rappresentativo di 51 quasar selezionati per le loro proprietà ottiche e radio. Evidentemente l'osservazione nello spettro X seleziona quasar con una luminosità ottica minore di quella dei quasar individuati con i metodi precedenti; la differenza di luminosità tra i due gruppi è il dato fondamentale per valutare con precisione l'emissione X di un quasar tipico e, di conseguenza, per determinare il contributo dei quasar al fondo a raggi X diffuso.

Anche se i quasar scoperti fortunosamente forniscono nuovi preziosi elementi per la comprensione del contributo che essi danno alla radiazione di fondo, l'esatta interpretazione di questi elementi è evidentemente molto complessa e per ora ambigua; come spesso accade nell'astronomia osservativa un grande passo avanti nella sensibilità degli strumenti ha portato nuove risposte, ma soprattutto nuove domande: esistono procedimenti di altro tipo che possano aiutare a distinguere il contributo dei quasar al fondo a raggi X da quello del plasma intergalattico? Una delle possibilità consiste nella determinazione di tutto lo spettro della radiazione osservata, cioè di quanta radiazione si rileva a ogni lunghezza d'onda rispetto a quella rilevata alle lunghezze d'onda vicine. Questa tecnica è, almeno in teoria, molto potente, perché spesso i diversi processi fisici fondamentali che producono raggi X danno luogo a spettri distinti e ben riconoscibili.

La spettroscopia alle lunghezze d'onda della radiazione X richiede tecniche sperimentali più difficili di quelle necessarie alle lunghezze d'onda del visibile; ciononostante diversi strumenti installati a bordo di satelliti sono riusciti negli ultimi anni a determinare con sempre maggiore precisione lo spettro del fondo a raggi X. Gran parte del lavoro in questo senso è stata compiuta da un gruppo presso il Goddard Space Flight Center della National Aeronautics and Space Administration; questi ricercatori hanno fatto notare che lo spettro osservato è molto simile a quello previsto per la radiazione di frenamento di un plasma caldo a una temperatura di circa 500 milioni di kelvin. Un risultato del genere si può interpretare come un notevole punto a favore della teoria secondo cui la radiazione di fondo deriva principalmente da un plasma caldo, mentre il contributo dei quasar (e di qualsiasi altro tipo di sorgente puntiforme lontana) è molto piccolo; come al solito, però, vi sono alcune complicazioni nell'interpretazione del risultato, che indeboliscono la validità della conclusione.

Una di queste difficoltà consiste nel capire che tipo di spettro è ragionevole attendersi se il fondo di radiazione proviene da un insieme di quasar invece che da un gas caldo. In questo caso lo spettro sarebbe la sovrapposizione dei singoli spettri di un gran numero di emittenti, ma questi spettri X sono stati determinati solo per pochi quasar tra i più luminosi e non risultano particolarmente simili a quello della radiazione di fondo. È possibile però che, sempre a causa degli effetti evolutivi, gli spettri dei quasar lontani siano diversi da quelli dei più vicini. Inoltre le determinazioni più accurate dello spettro del fondo diffuso riguardano lunghezze d'onda leggermente diverse da quelle delle immagini dell'Osservatorio Einstein: se la natura fosse talmente malevola da utilizzare in prevalenza meccanismi di emissione diversi nelle due bande, coloro che studiano le caratteristiche dello spettro del fondo di radiazione senza saperlo potrebbero avere di fronte un problema leggermente diverso da quello affrontato dai ricercatori che si occupano delle osservazioni fotografiche. Sarebbe però possibile una misurazione abbastan-

za univoca che aiuterebbe a distinguere il contributo al fondo di radiazione dovuto alle fonti discrete lontane da quello di un plasma caldo; nelle riprese dell'Osservatorio Einstein, come ho già ricordato, il numero delle sorgenti fortunose rivelate cresceva all'aumentare del tempo di esposizione dell'immagine; pertanto il contributo di queste emittenti discrete al totale dei raggi X registrati nell'immagine è sicuramente una stima per difetto, molto prudente, della frazione della radiazione dovuta a tutte le sorgenti discrete. Nelle esposizioni più prolungate, questa frazione arriva a un quarto; è chiaro perciò che oggetti come i quasar offrono un contributo non indifferente alla radiazione di fondo. Questo valore tuttavia rimane ben al di sotto del 50 per cento necessario per poter affermare con una certa fiducia che la maggior parte del fondo di radiazione è dovuta a oggetti discreti e con questo risolvere il problema.

Dalla discussione precedente dovrebbe risultare chiaro che uno strumento, in grado di formare immagini X, più potente di quelli a bordo dell'Osservatorio Einstein potrebbe risolvere la questione in modo immediato ed elegante, e cioè semplicemente risolvendo individualmente abbastanza sorgenti deboli da spiegare almeno metà dell'emissione totale di ogni area del cielo, oppure scoprendo che, anche a livelli molto bassi di luminosità X, la radiazione rimane effettivamente diffusa e non si scinde in sorgenti separate. Le tecnologie richieste sono già disponibili e la NASA ha già ricevuto una proposta per uno strumento di potenza appropriata da installare su un satellite artificiale; lo sviluppo di questo progetto, noto come AXAF (Advanced X-ray Astrophysics Facility) richiederebbe circa 10 anni dallo stanziamento dei fondi. Un comitato di astronomi presieduto da George B. Field della Harvard University ha elaborato un elenco, in ordine di priorità, di missioni interessanti per il prossimo decennio, e l'AXAF era al primo posto.

Non è ben chiaro attualmente se uno studio più approfondito dei quasar a raggi X scoperti fortunosamente potrà metterci in grado di capire con una precisione ragionevole gli eventuali effetti dell'evoluzione dei quasar; se così fosse, però, questa comprensione, insieme con dati più raffinati sul numero di quasar presenti nelle varie zone del cielo, potrebbe portare a un calcolo preciso dell'importanza dei quasar per il fondo di radiazione X e, per conseguenza, a una valutazione del contributo del plasma intergalattico caldo. I dati osservativi, particolarmente precisi, che occorrono saranno quasi certamente forniti dallo Space Telescope, la cui entrata in orbita è prevista per il 1985; in quel momento, o eventualmente con il lancio del satellite AXAF all'inizio degli anni novanta, sarà forse possibile finalmente dire con una certa sicurezza se il fondo a raggi X diffuso e i quasar, scoperti indipendentemente 20 anni fa, sono effettivamente due aspetti dello stesso fenomeno.

Macchine che camminano

I microelaboratori hanno consentito il controllo di macchine dotate di gambe: esperimenti eseguiti con tali dispositivi ci fanno capire meglio la dinamica del cammino e della corsa

di Marc H. Raibert e Ivan E. Sutherland

Numerose sono le macchine che imitano la natura: per fare un esempio, l'aeroplano imita, almeno in parte, l'uccello in volo. Fino a oggi, però, il camminare, una forma di locomozione animale che ci è ben nota, non ha trovato imitazioni nel mondo delle macchine. Ci siamo chiesti se i moderni calcolatori e i sistemi di controllo a retroazione avrebbero consentito di costruire macchine capaci di camminare e abbiamo svolto uno studio in tal senso avvalendoci sia di modelli al calcolatore sia di alcuni prototipi appositamente costruiti.

Finora abbiamo realizzato due macchine; la prima, dotata di sei gambe e di conduzione, serve per studiare la forma di locomozione tipica degli insetti, per la quale non occorre tener conto dell'equilibrio. L'altra ha una sola gamba, si muove a salti e ci è utile per studiare i problemi connessi con l'equilibrio. Abbiamo chiamato andatura radente il primo tipo di locomozione per distinguerlo dall'andatura alta, in cui entrano in gioco capacità di equilibrio, e dalla corsa, che comprende anche periodi di sollevamento, cioè fasi in cui tutti gli arti perdono il contatto con il terreno.

A differenza della ruota, che modifica con continuità e con gradualità il suo punto di appoggio sorreggendo sempre il peso di cui è gravata, una gamba cambia il suo punto d'appoggio sul terreno in maniera discreta, discontinua, e perché ciò sia possibile, deve prima poter scaricare tutto il peso che su di essa grava. In altre parole, affinché un sistema dotato di gambe possa camminare o correre, è necessario che per ogni gamba vi siano alternativamente periodi nei quali è soggetta a un carico, mentre il relativo piede appoggia sul terreno, e periodi in cui non deve sostenere peso e di conseguenza il piede è libero di muoversi. Questo tipo di avvicendamento ciclico fra una fase con carico, detta di appoggio, e una fase senza carico, detta di oscillazione o di sospensione, si ritrova in qualsiasi sistema dotato di gambe. Come ben sa chiunque abbia montato un cavallo al trotto o al galoppo l'avvicendamento fra fase di appoggio e fase di sospensione può generare un pronunciato ondeggiamento verticale («su e giù»).

Il lavoro del nostro e di altri gruppi può condurre alla realizzazione di macchine in grado di camminare con andatura radente e alta o di correre su terreni tanto morbidi o tanto accidentati da rendere inutilizzabili i veicoli a ruote: si tratta quindi di una ricerca passibile di applicazioni in campo industriale, agricolo e militare. Il vantaggio dei veicoli dotati di gambe sui terreni difficili è rappresentato dal fatto che possono scegliere punti di appoggio per aumentare la trazione, per minimizzare gli sbandamenti e per superare ostacoli. In linea di principio le prestazioni dei veicoli dotati di gambe sono largamente indipendenti dalla scabrosità del terreno. Il nostro obiettivo è stato quello di valutare i calcoli necessari per controllare e coordinare i movimenti delle gambe. È chiaro che fra i componenti più importanti di macchine capaci di camminare o di correre con regolarità dovranno figurare raffinatissimi programmi di controllo.

Come è stato già detto, la locomozione può avvenire con o senza equilibrio dinamico. Gli animali che camminano radenti al suolo non hanno il problema dell'equilibrio, in quanto posseggono sei o più zampe, almeno la metà delle quali possono essere impiegate per assicurare con continuità l'appoggio su tre punti. Riprese cinematografiche ad alta velocità mostrano infatti che gli insetti hanno un'andatura in cui gli arti si alternano a tre a tre.

Anche se è possibile realizzare una macchina ad andatura radente, e che quindi non necessita di un controllo dinamico dell'equilibrio, con quattro gambe sole, essa procederebbe però in modo goffo perché a ogni passo il peso deve essere scaricato su arti diversi, per impedire che la macchina si ribalti. Buone prestazioni senza un sistema di equilibrio attivo richiedono almeno sei gambe; è questo il numero minimo che garantisce con continuità l'appoggio su tre punti anche quando metà delle gambe è sollevata. Finora si sono costruite diverse macchine a sei gambe, che differiscono per dimensioni e progetto meccanico, ma che impiegano tutte un controllo delle gambe mediante calcolatore.

Un programma di calcolatore che controlli una macchina del genere esegue cinque funzioni. In primo luogo regola l'andatura della macchina, ossia la successione e il modo in cui le gambe assolvono al compito della locomozione. Le macchine a sei gambe hanno andature in cui le fasi di oscillazione possono interessare una gamba per volta oppure due o tre gambe contemporaneamente.

Le andature, le più semplici delle quali comportano una successione regolare nei movimenti delle gambe, possono essere descritte con un'opportuna notazione della successione: così per esempio un'andatura su tre piedi alternati può essere indicata con (1,5,3; 6,4,2), ove le virgole denotano l'uso contemporaneo delle gambe e il punto e virgola l'uso in successione, mentre un'andatura le cui fasi di oscillazione interessano un solo piede per volta può essere rappresentata come (3;2;1;4;5;6;) oppure (3;4;2;5;1;6;). Un'andatura con diverse gambe sollevate contemporaneamente rende in genere possibile un avanzamento più rapido, ma assicura una stabilità minore rispetto a una in cui sia maggiore il numero delle gambe al suolo nella fase di appoggio.

Una seconda funzione assegnata a un programma di calcolatore per il controllo di una macchina ad andatura radente è di evitarne il ribaltamento. Se il baricentro viene a trovarsi all'esterno dell'area delimitata dalle gambe, la macchina si capovolge. Il calcolatore deve verificare la posizione del baricentro rispetto a quella delle gambe per assicurare con continuità una base di appoggio sufficientemente ampia. Con le andature semplici la geometria delle gambe può essere di per se stessa sufficiente a fornire una base di appoggio adeguata, ma nel caso di andature più complesse assume un'importanza critica il calcolo della stabilità statica.

Una terza funzione del calcolatore di controllo è la distribuzione del carico e delle forze laterali fra le varie gambe su cui viene ripartito il peso della macchina. Nell'andatura a tre gambe alternate, per esempio, la distribuzione del carico è determinata dalla geometria delle tre gambe di sostegno. Con più di tre gambe di sostegno, invece, il calcolatore deve decidere come effettuare la distribuzione delle sollecitazioni per raggiungere obiettivi di livello superiore, quali la scorrevolezza del procedere e un'andatura in cui siano minimi i disturbi indotti dal suolo.

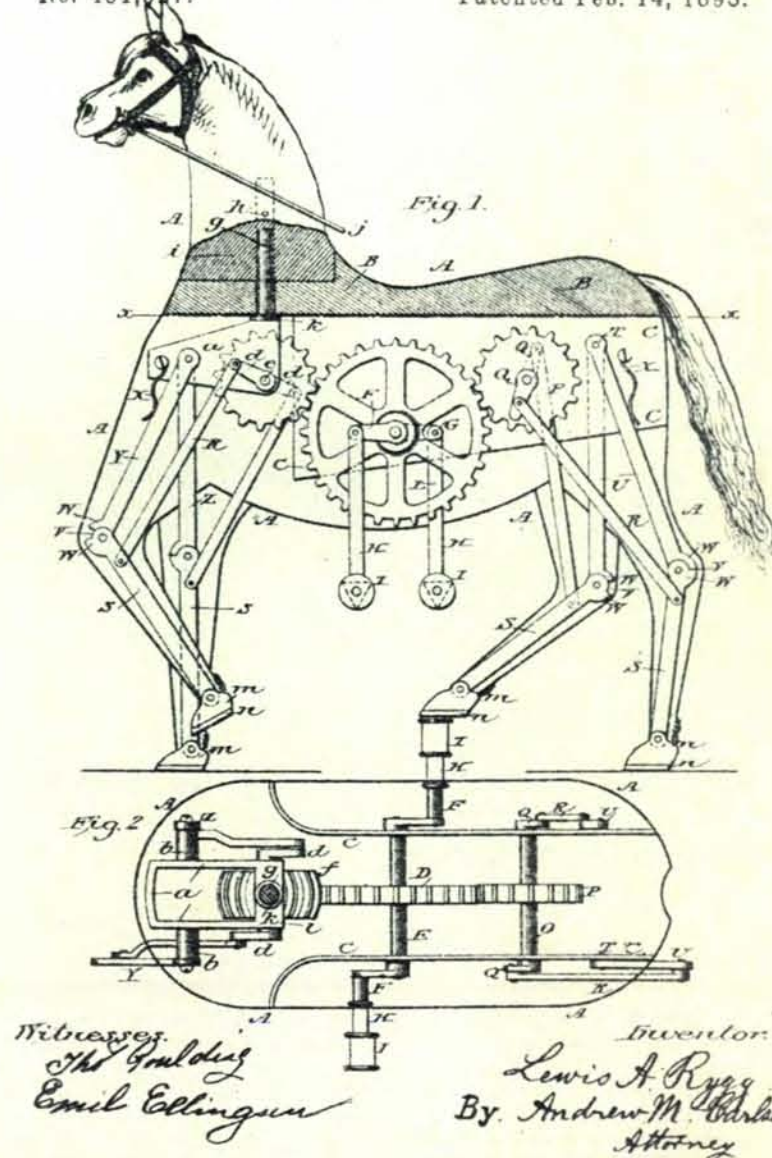
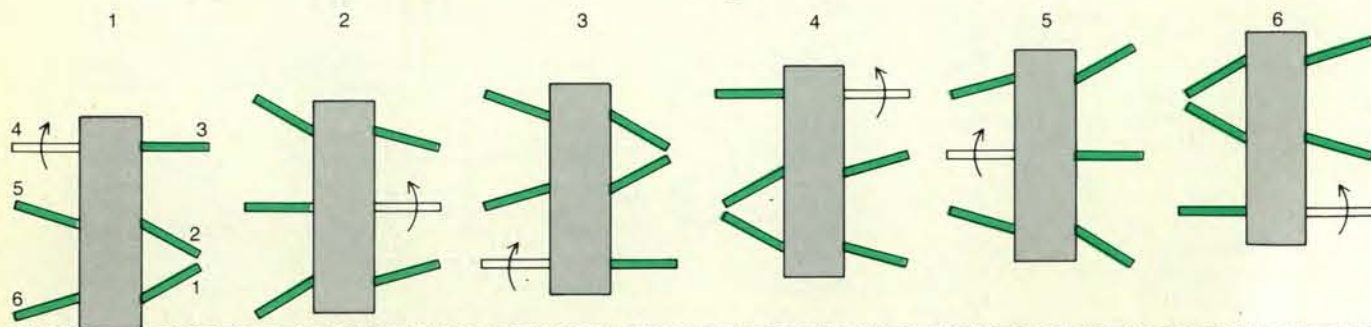
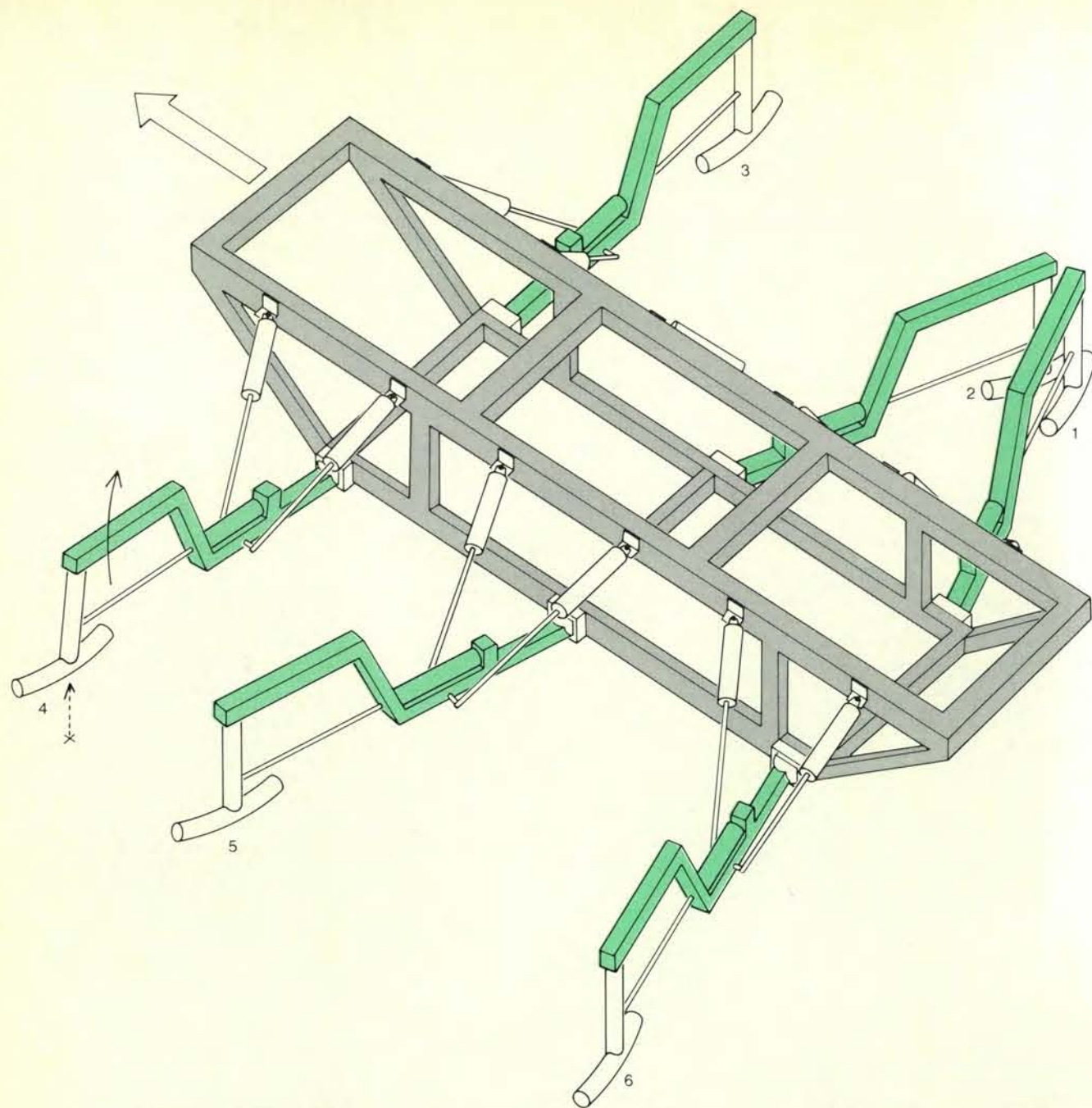
Come accennato, il programma di controllo deve provvedere a distribuire le sollecitazioni laterali anche quando solo tre gambe sorreggono la macchina. Il problema può essere visto in questo modo: il sistema di controllo deve evitare che la macchina compia semplicemente esercizi isometrici contro il suolo. La quantità di informazioni che debbono essere raccolte dai sensori e la quantità di calcoli richiesta per distribuire i carichi laterali fra molte gambe possono essere molto grandi. Per la macchina ad andatura radente in fase di costruzione abbiamo ridotto la complessità del compito prevedendo circuiti idraulici passivi che distribuiscano in modo automatico le sollecitazioni laterali.

La quarta funzione del calcolatore di controllo è di fare in modo che le gambe non siano costrette a superare una certa ampiezza di movimento. La geometria delle gambe può essere tale da rendere possibile un urto occasionale fra due gambe contigue: se l'interferenza è possibile, il calcolatore deve limitare i movimenti delle gambe per prevenire eventuali danni. Per ottimizzare l'utilità di ciascun arto il posizionamento della gamba sul suolo deve tener conto sia dei suoi limiti di movimento, sia del movimento che la macchina può compiere durante la sua fase di appoggio. Così, per esempio, se la macchina deve voltare a destra, le gambe anteriori dovranno essere posizionate più a destra dell'usuale affinché sia possibile coordinare il loro spostamento laterale durante la deviazione. Nel caso di un veicolo dotato di controllo autonomo il posizionamento delle gambe può essere basato sul previsto percorso futuro del veicolo, mentre per i veicoli con conducente il corretto posizionamento di ogni gamba richiede che il conducente preveda i comandi da impartire durante la successiva fase di appoggio.

La quinta funzione del calcolatore di controllo è la scelta dei punti in grado di assicurare un supporto adeguato. Su terreno regolare si tratta di un compito semplice da assolvere, che però su terreni accidentati può diventare straordinariamente complesso. Non è stato ancora costruito un sistema in grado di assolvere a questa funzione. Si potrebbe concepire un sistema di scansione che esamini il terreno davanti alla macchina e che scelga, sulla base delle informazioni raccolte, i possibili punti di appoggio. Per utilizzare un simile dispositivo, il calcolatore di controllo dovrebbe costruire un modello numerico interno del terreno, che tenga conto solo delle irregolarità di dimensioni maggiori o uguali a quelle del piede della macchina. Se un operatore umano fosse in grado di fornire ulteriori informazioni al modello, la valutazione dei possibili punti di appoggio potrebbe essere facilitata.

Uno di noi (Sutherland) sta realizzando una macchina ad andatura radente a trasmissione idraulica azionata da un motore a benzina; le sei gambe, mosse da attuatori idraulici, non richiedono un controllo dell'equilibrio dinamico.

Un microelaboratore incorporato controlla le gambe mediante l'apertura e la chiusura delle valvole che regolano il flusso dell'olio agli attuatori. Ciascuna gamba contiene alcuni sensori che trasmettono al microelaboratore informazioni sulla posizione e sulle forze agenti sulla gamba stessa. La macchina è di dimensioni sufficienti per ospitare un conducente, che ne controlla la velocità e la direzione e ne stabilisce l'inclinazione e la distanza dal terreno. La velocità del veicolo è di circa 3,2 chilometri all'ora.

Un obiettivo dello studio era di minimizzare le quantità di calcoli necessari per ottenere una andatura radente. I circuiti idraulici sono progettati in modo da far muovere le gambe lungo percorsi utili senza l'intervento del microelaboratore, che si limita semplicemente a scegliere per ciascuna gamba uno dei percorsi disponibili nel programma. Così il microelaboratore è libero di dedicarsi alla scelta delle gambe da destinare alla funzione di sostegno e di decidere dove dirigere il passo successivo, senza impiegare tempo e capacità di calcolo per stabilire nei dettagli i movimenti delle gambe stesse.

Lewis A. Rygg ottenne nel 1893 il brevetto per questo cavallo meccanico. Il disegno in basso è una vista in sezione lungo la linea *x-x* della *Fig. 1*. Le staffe funzionavano come pedali per mettere in grado il cavaliere di imprimere il movimento. La direzione veniva impartita con redini che muovevano la testa e le zampe anteriori. La macchina, forse mai costruita, sarebbe stata simile a molti moderni giocattoli che camminano, che non disponendo di sensori né di dispositivi di calcolo, non possono seguire l'andamento del terreno, ma sono in grado solo di avanzare su superfici piane.

La macchina a sei gambe costruita da uno degli autori (Sutherland) ha una andatura radente e non richiede controlli di equilibrio. Le sei gambe sono controllate da un microcalcolatore incorporato e mosse da un motore a benzina da 18 cavalli-vapore che aziona pompe idrauliche separate per le gambe di destra e di sinistra. Un conducente guida la macchina facendo fluire l'olio a velocità differenti nei due circuiti laterali. Appositi sensori comunicano al microcalcolatore i comandi del conducente, le posizioni di ciascuna gamba e le forze che agiscono su di esse; in base alle informazioni ricevute, il microcalcolatore stabilisce l'ordine e il percorso del moto delle gambe. Sei gambe assicurano la stabilità, perché almeno tre appoggiano sempre al terreno. Circuiti idraulici passivi semplificano le mansioni di calcolo; una gamba che sorregge del peso può essere o connessa all'unità di comando o spinta lungo il terreno secondo il movimento generale delle altre gambe. Il diagramma in basso indica la posizione delle gambe nel ciclo di spostamento che può essere indicato con (4;2;6;3;5;1;). Un rettangolo in colore rappresenta una gamba che tocca il suolo e un rettangolo bianco una gamba alzata che si sposta in avanti come indicato dalla freccia. Nella rappresentazione numerica del ciclo di spostamento i punti e virgola indicano l'uso sequenziale delle sei gambe.

Ogni gamba può muoversi in avanti o all'indietro e in alto o in basso, grazie all'articolazione dell'«anca» che la collega al telaio della macchina. L'esecuzione dei movimenti avviene con l'estensione o la ritrazione di due attuatori idraulici disposti a *V* al disopra di ciascuna gamba. Una posizione delle valvole prevede che l'olio liberato da un attuatore sia inviato nell'altro, cosicché, se l'uno si ritrae, l'altro si estende nella stessa misura. A causa della geometria degli assi di rotazione il risultato è un movimento orizzontale della gamba.

Il movimento orizzontale può essere azionato o meno dal motore, a seconda della posizione delle valvole, e quindi si può fare in modo che alcune gambe contribuiscano all'avanzamento della macchina e altre seguano solamente. Le gambe, una volta appoggiate a terra e sottoposte al carico, possono muoversi liberamente in avanti o all'indietro per seguire il movimento delle gambe «motrici». Perciò il calcolatore di controllo, per ottenere un avanzamento regolare, non deve calcolare l'istante preciso in cui la gamba tocca il suolo e i dettagli del movimento necessario al momento del contatto.

L'articolazione del «ginocchio» di ogni gamba è azionata da un attuatore idraulico separato, disposto orizzontalmente, che può essere messo in funzione quando la gamba è sollevata, per spostare lateralmente il piede durante il passo. Quando il piede appoggia al terreno, questa articolazione deve potersi muovere leggermente per adattare il percorso circolare del ginocchio, rispetto all'articolazione dell'anca, al percorso rettilineo del piede sul terreno. Si tratta di un movimento complesso che non richiede però l'intervento del calcolatore, in quanto è sufficiente un semplice collegamento in parallelo di tali attuatori per permettere a tutti i ginocchi di adattarsi al moto del veicolo. Una pompa idraulica supplementare presente nel sistema consente di effettuare un movimento laterale di tutte queste ultime articolazioni facendo spostare la macchina lateralmente come un granchio.

Con la macchina in grado di saltare in due dimensioni è possibile studiare il problema del controllo del moto quando è in gioco anche l'equilibrio. La macchina funziona di norma disposta parallelamente a una parete inclinata da cui è separata da un cuscino d'aria; essa può inclinarsi solamente nel piano definito dalla parete. Un calcolatore riceve i dati dei sensori d'angolo, da un sensore di pressione e da un interruttore disposto nel piede; il calcolatore controlla i salti regolando la pressione nella camera pneumatica che funziona come una molla. Il calcolatore applica anche una coppia di forze fra la gamba e il corpo per regolare l'angolo dell'articolazione dell'anca, che determina lo spostamento orizzontale del piede e di conseguenza influenza l'equilibrio.

Il conducente della macchina può intervenire su tre tipi di controlli. In primo luogo può regolare il flusso dell'olio, controllando lo spostamento delle pompe idrauliche. La macchina è fornita di pompe separate per le gambe di destra e di sinistra onde consentire il cambiamento di direzione facendo avanzare la macchina più velocemente da un lato che dall'altro. Le regolazioni dei controlli di direzione sono trasmesse al microelaboratore affinché questo possa posizionare i piedi in modo opportuno. Se per esempio la macchina deve dirigersi a destra, per seguire la rotazione il piede anteriore deve essere trasferito verso destra e quello posteriore verso sinistra. Se la macchina deve spostarsi all'indietro (il che si ottiene invertendo il flusso dell'olio), a ogni passo il piede deve essere trasferito all'indietro anziché in avanti. Al momento in cui un piede viene sollevato da terra, il calcolatore di controllo sceglie la posizione che esso deve raggiungere, in funzione della portata e della direzione del flusso stabilite dal conducente. Quando un piede si avvicina al limite di spostamento consentito, il calcolatore di controllo ne inizia il sollevamento per trasferirlo su un nuovo punto di appoggio. Se poi un piede raggiunge tale limite, il microelaboratore arresta il veicolo fino a quando quel piede può essere sollevato da terra e trasferito a un nuovo punto di appoggio dove abbia spazio per muoversi.

Il secondo tipo di controllo a disposizione del conducente è quello dell'assetto e della distanza della macchina dal terreno. Il conducente può agire su un comando che modifica la posizione verticale di supporto dei piedi di destra e di quelli di sinistra, facendo inclinare il veicolo di conseguenza da un lato. In modo analogo è possibile modificare la posizione verticale di supporto dei piedi anteriori o posteriori per inclinare la macchina in avanti o all'indietro. Un altro comando consente di estendere o ritrarre contemporaneamente le sei gambe per regolare la distanza del veicolo dal terreno.

Il terzo tipo di controllo consente al conducente di posizionare con cura i piedi della macchina per l'avanzamento su terreni molto accidentati. Non abbiamo però ancora deciso come effettuare questo controllo. Un carro semovente costruito alcuni anni fa da Ralph Moser presso la General Electric Company era basato su un controllo solo manuale della posizione dei piedi ed era quindi faticosissimo da guidare. Riteniamo che anche la scelta dell'andatura possa essere importante, ma non disponiamo ancora di sufficiente esperienza per stabilire se questa scelta potrà essere automatica o se sarà necessario un comando manuale. Proprio per rispondere a interrogativi di questo genere abbiamo costruito la macchina.

Altri temi dei quali ci siamo occupati sono il cammino e la corsa, in cui l'equilibrio ha un ruolo importante. Fino a un secolo fa veniva dibattuta la questione se un cavallo al trotto alzasse contemporaneamente o meno tutte e quattro le zampe dal suolo. Le fotografie «istantanee» di Eadweard Muybridge posero fine alla questione dimostrando che un cavallo al trotto alza contemporaneamente le quattro zampe dal suolo. Una persona che

corre si comporta nello stesso modo e lo stesso si può dire di altri animali come il cane, il ghepardo e ovviamente il canguro. Tali animali quindi non solo sono in grado di camminare, il che richiede un equilibrio dinamico, ma anche di correre sfruttando con efficacia un tipo di moto balistico per aumentare la velocità di avanzamento.

Due sono le differenze fondamentali fra un veicolo a equilibrio statico e un veicolo a equilibrio dinamico. La prima differenza risiede nella definizione di stabilità. Un veicolo ad andatura radente è stabile se le sue gambe forniscono sempre almeno tre punti di appoggio, in modo che non si ribalti; un veicolo dotato di equilibrio dinamico capace di camminare o correre può inclinarsi per brevi intervalli di tempo senza inconvenienti. Grazie ai movimenti delle gambe e del corpo, ogni singolo intervallo in cui il veicolo si inclina è di breve durata e in media viene mantenuta una base di appoggio sufficiente. Per esempio, un uomo che corre appoggia al terreno alternativamente le due gambe, e di conseguenza il corpo ha una base di appoggio solo a intervalli.

La seconda differenza fra equilibrio statico ed equilibrio dinamico riguarda la velocità e la quantità di moto. L'equilibrio statico presuppone che la configurazione delle gambe di sostegno e la posizione del baricentro siano sufficienti a determinare la stabilità, indipendentemente dal movimento del veicolo. Non sempre però queste valutazioni di statica sono sufficienti: un veicolo che si muove velocemente può per esempio ribaltarsi in avanti se arrestato bruscamente e se il baricentro è troppo vicino alle gambe anteriori. Per capire la maggiore mobilità dei sistemi capaci di camminare e di correre, occorre modificare la definizione di stabilità e tener conto della velocità nel calcolare l'equilibrio.

Allo scopo di studiare il problema dell'equilibrio nella sua forma più semplice, uno di noi (Raibert) ha costruito e fatto funzionare, con i suoi collaboratori alla Carnegie Mellon University, una macchina con una sola gamba che salta e corre come un canguro, con una successione di balzi. Il dispositivo può essere considerato un trampolo a molla controllato dal calcolatore. Lo studio è stato incoraggiato dalla considerevole semplicità dell'algoritmo di equilibrio. Il movimento della macchina nella configurazione attuale può avvenire solo in un piano e quindi essa può inclinarsi in una sola direzione.

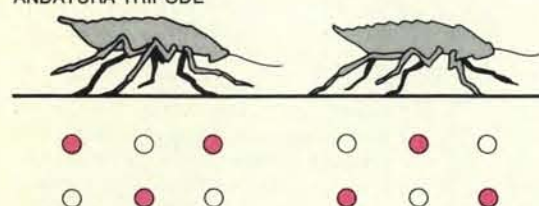
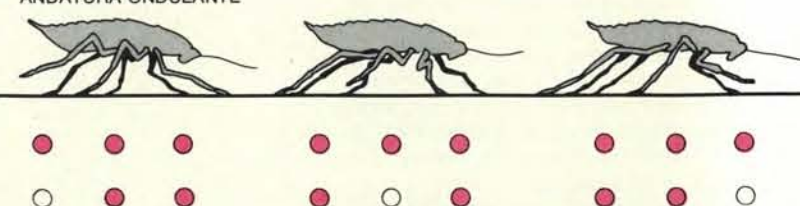
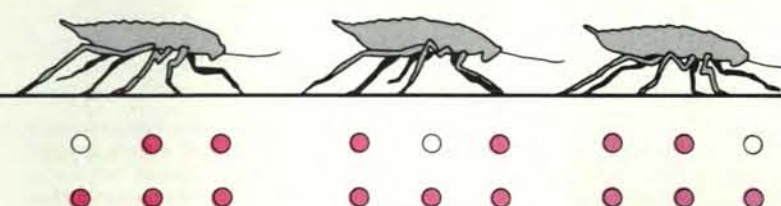
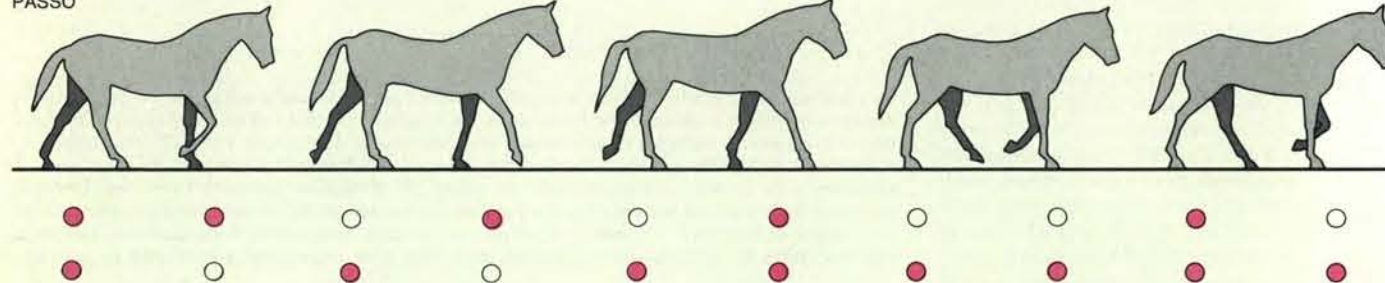
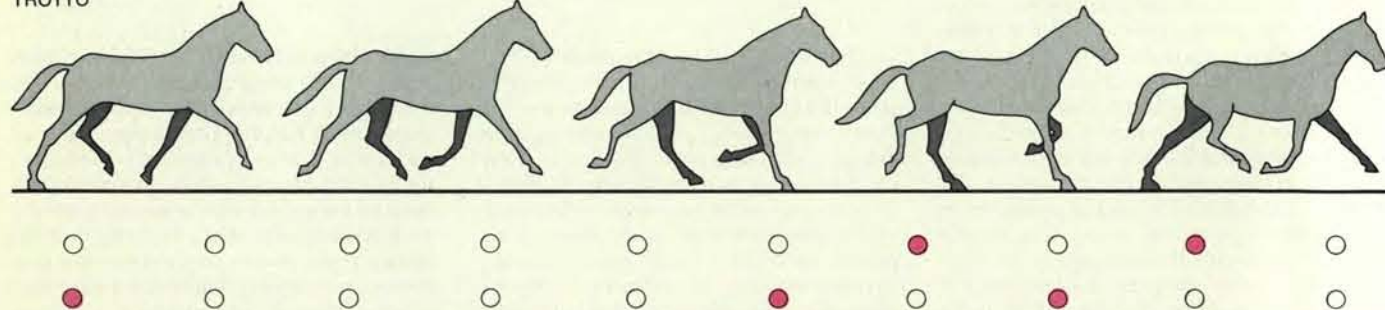
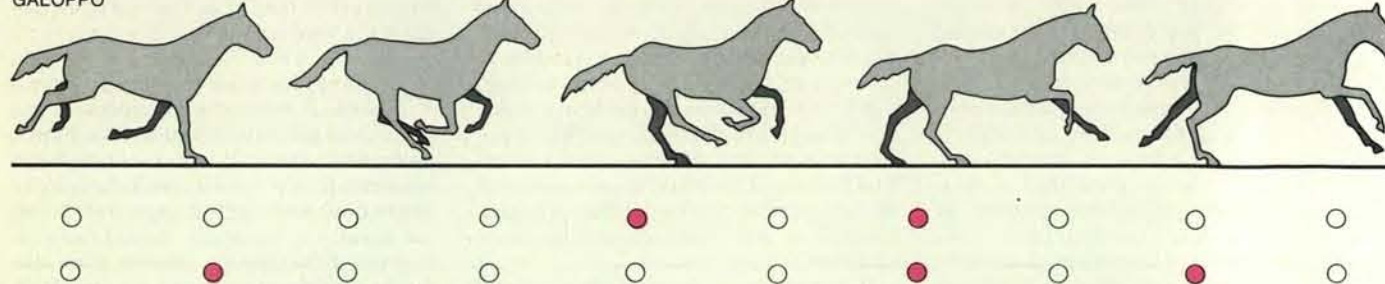
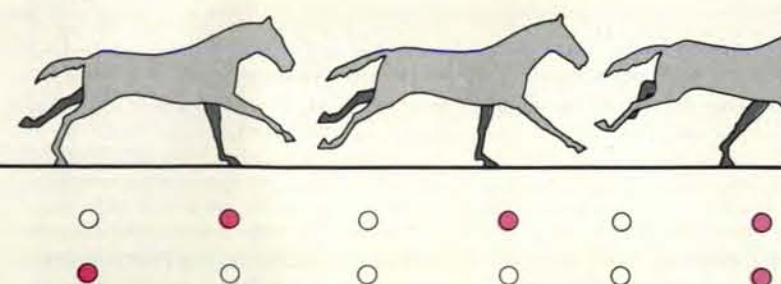
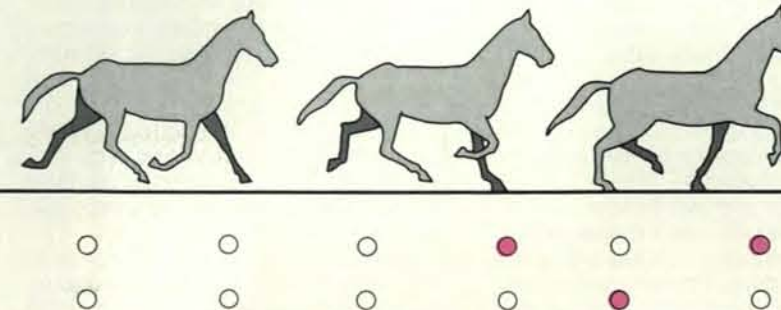
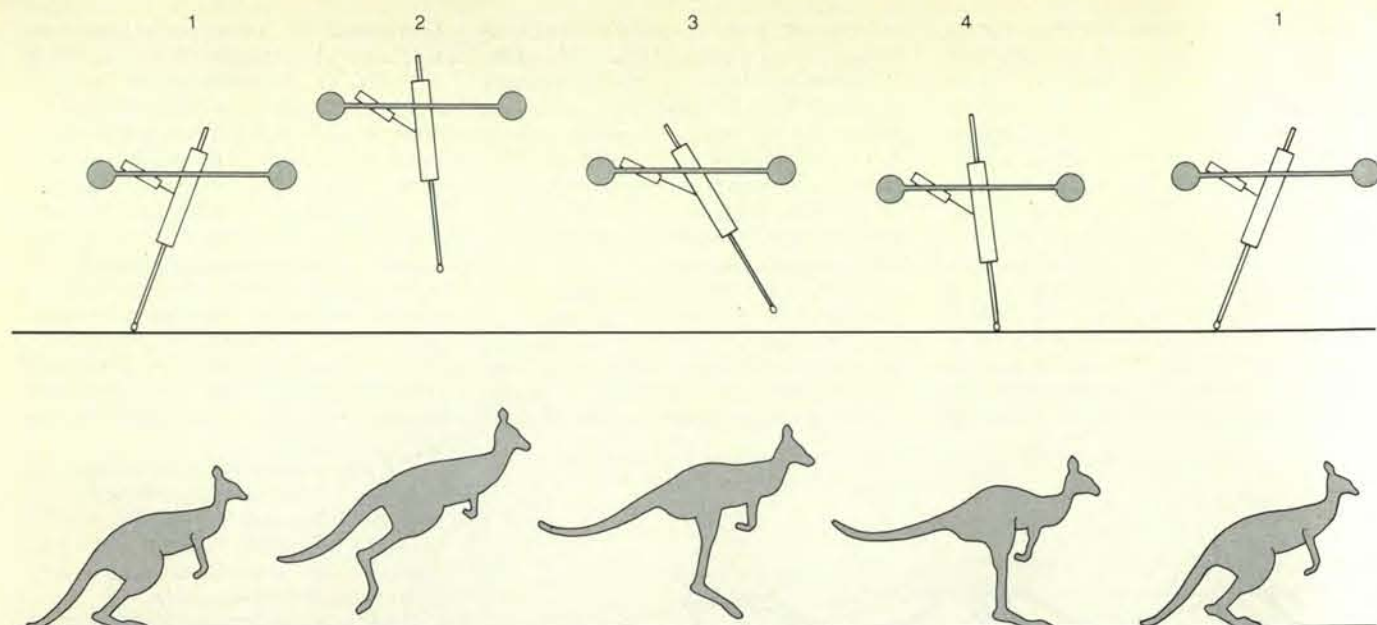
La macchina è composta da due parti principali: un corpo e una gamba. Il corpo costituisce la struttura principale e alloggia le valvole, i sensori e i circuiti elettronici. La gamba è un meccanismo semplice, di lunghezza variabile nel senso del suo asse, in grado di ruotare rispetto al corpo intorno a una cerniera detta «articolazione dell'anca». La gamba appoggia su una molla, la cui tensione è regolabile, in modo molto simile a una gamba umana con i suoi muscoli e tendini elastici. La molla è costituita da un cilindro ad aria, le cui pressioni vengono controllate da sensori e valvole. In fondo alla gamba è montato un piccolo piede.

Il movimento oscillante della gamba è controllato da un secondo attuatore ad aria, che applica una coppia di forze alla cerniera. Una semplice valvola a due posizioni (aperto-chiuso) controlla la molla; l'angolo di oscillazione richiede un controllo mediante una servo-valvola proporzionale, cioè un dispositivo a retroazione la cui uscita sia proporzionale alla intensità del segnale di ingresso. Poiché il momento d'inerzia della gamba è inferiore al 10 per cento del momento d'inerzia del corpo, la gamba può oscillare quando è sollevata senza impartire forti movimenti al corpo. L'inclinazione del corpo è misurata da un giroscopio, che consente al calcolatore di controllo di mantenere il corpo stesso in un assetto di equilibrio. Altri sensori misurano l'angolo dell'articolazione, la lunghezza della gamba, la pressione nel cilindro che contiene la molla, l'angolo fra la gamba e il terreno e la forza con cui la gamba appoggia sul terreno.

Tre differenti circuiti di retroazione regolano il funzionamento della macchina: uno il movimento verticale, uno l'equilibrio e l'ultimo l'assetto del corpo. Ogni circuito è sincronizzato con il movimento fondamentale a balzi.

Il primo circuito di retroazione controlla l'altezza del balzo incrementando o limitando, a ogni salto, l'energia fornita per raggiungere l'altezza voluta e compensare l'energia perduta. Il circuito esegue entrambi i compiti con l'iniezione o l'aspirazione di aria dal cilindro principale di azionamento per regolare la tensione effettiva della molla ad aria. In altre parole il controllo di altezza comanda la sincronizzazione e la quantità di potenza erogata al meccanismo di avanzamento, ottenendo il raggiungimento dell'altezza voluta per il balzo. Raggiunta questa condizione, la maggior parte dell'energia necessaria per il balzo seguente viene recuperata dalla molla nella quale era stata immagazzinata durante l'atterraggio precedente. Fin quando i balzi sono abbastanza uniformi, il compito di regolare l'energia necessaria per effettuarli non è particolarmente complesso.

Il secondo circuito di retroazione assicura l'equilibrio della macchina determinando la posizione del piede quando è sollevato, perché l'atterraggio possa portare a una posizione di appoggio in equilibrio. Il calcolo della corretta posizione del piede tiene conto della velocità di avanzamento del veicolo e dell'inclinazione del corpo. L'equilibrio è assicurato sempre dal medesimo algoritmo, indipendentemente dal fatto che la macchina salti sul posto, acceleri per correre, corra a velocità costante, sorpassi oggetti o rallenti per tornare a saltare sul posto.

Quando la macchina salta sul posto, la gamba e il piede compiono piccoli movimenti per compensare l'azione di disturbi esterni e gli errori dei salti precedenti. Quando la macchina inizia a correre, per esempio verso destra, il piede viene mosso inizialmente a sinistra per sbilanciare il veicolo affinché si inclini nella direzione voluta. Una corsa costante è molto simile ai salti sul posto, salvo per il fatto che le regolazioni di equilibrio integrano gli ampi movimenti della gamba, determinati dalla velocità di spostamento. L'arresto avviene in modo analogo alla partenza, salvo il fatto che la macchina si inclina nella direzione opposta a quella del movimento.

L'andatura degli insetti ha suggerito il principio fondamentale di progetto per il metodo di locomozione della macchina ad andatura radente a sei gambe, illustrata sulla copertina di questo fascicolo. Un cerchietto in colore indica l'appoggio, un cerchietto bianco il sollevamento delle zampe dal terreno. Un'andatura tripode assicura sempre la stabilità. Nel movimento con andatura ondulante le zampe adiacenti dell'insetto si muovono in successione.

Le andature di un cavallo rappresentano il tipo di locomozione in cui l'equilibrio ha una funzione importante. Nel passo almeno due sono appoggiate contemporaneamente al terreno; nel trotto e nel galoppo vi sono momenti in cui tutte e quattro le zampe sono staccate dal terreno. I disegni sono basati sulle riprese fotografiche a velocità variabile eseguite da Eadweard Muybridge 100 anni fa e che permisero di appurare questo fatto. Molti altri animali corrono nello stesso modo.

La macchina a una gamba funziona in modo ciclico, come tutti i sistemi dotati di gambe; la gamba alterna fasi di appoggio e fasi di sospensione. A sinistra la macchina sta per iniziare un balzo; mentre è in aria la gamba oscilla in avanti preparandosi all'atterraggio successivo, avvenuto il quale si ha la compressione della molla della gamba fino a raggiungere la minima estensione in preparazione del balzo successivo. Un sensore di contatto con il suolo agisce come innesco per il programma di controllo verticale. La macchina dispone anche di circuiti a retroazione per controllare l'assetto e l'equilibrio in sincronia con il controllo verticale. Come un trampolo a molla, la macchina può mantenersi in equilibrio solo quando salta e il balzo che compie è simile a quel tipo di movimento di un canguro che Muybridge chiamava *«ricochet»*.

Il terzo circuito di retroazione stabilizza l'assetto del corpo per mantenerlo eretto. Esso fornisce le coppie di forze fra la gamba e il corpo mentre il piede è appoggiato al terreno al fine di raggiungere l'assetto voluto durante la successiva fase di sospensione. L'efficacia dell'asservimento dipende dall'attrito fra piede e terreno. Il servomeccanismo di assetto che opera nella fase di appoggio agisce anche sul meccanismo di azionamento dell'articolazione impiegato dal circuito di retroazione che lavora quando la macchina è sollevata. Alcuni ingegnosi dettagli per passare da un modo di controllo all'altro sono connessi con il rilevamento dell'inizio e della fine di ogni fase di sollevamento. Il meccanismo che applica la coppia deve rimanere inattivo durante queste fasi, per evitare che il piede slitti sul terreno.

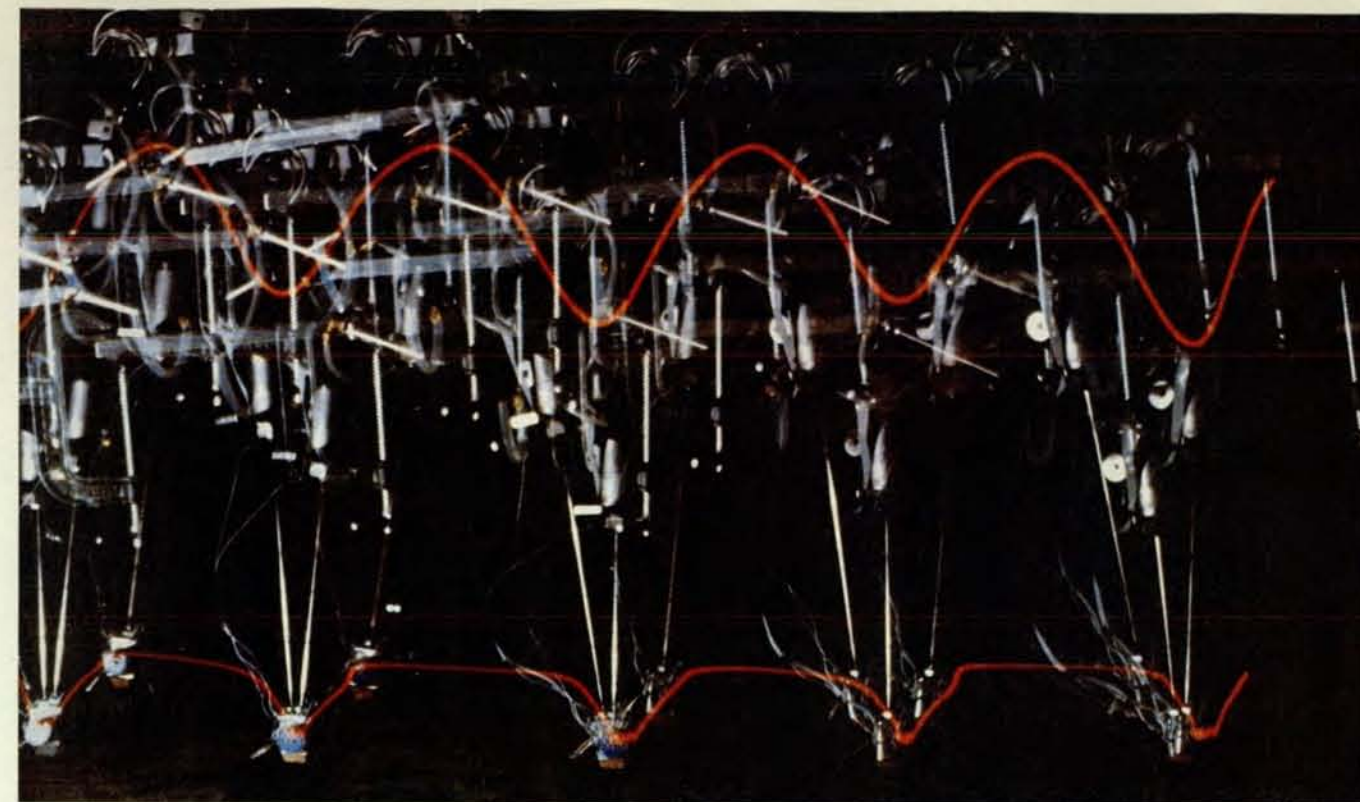
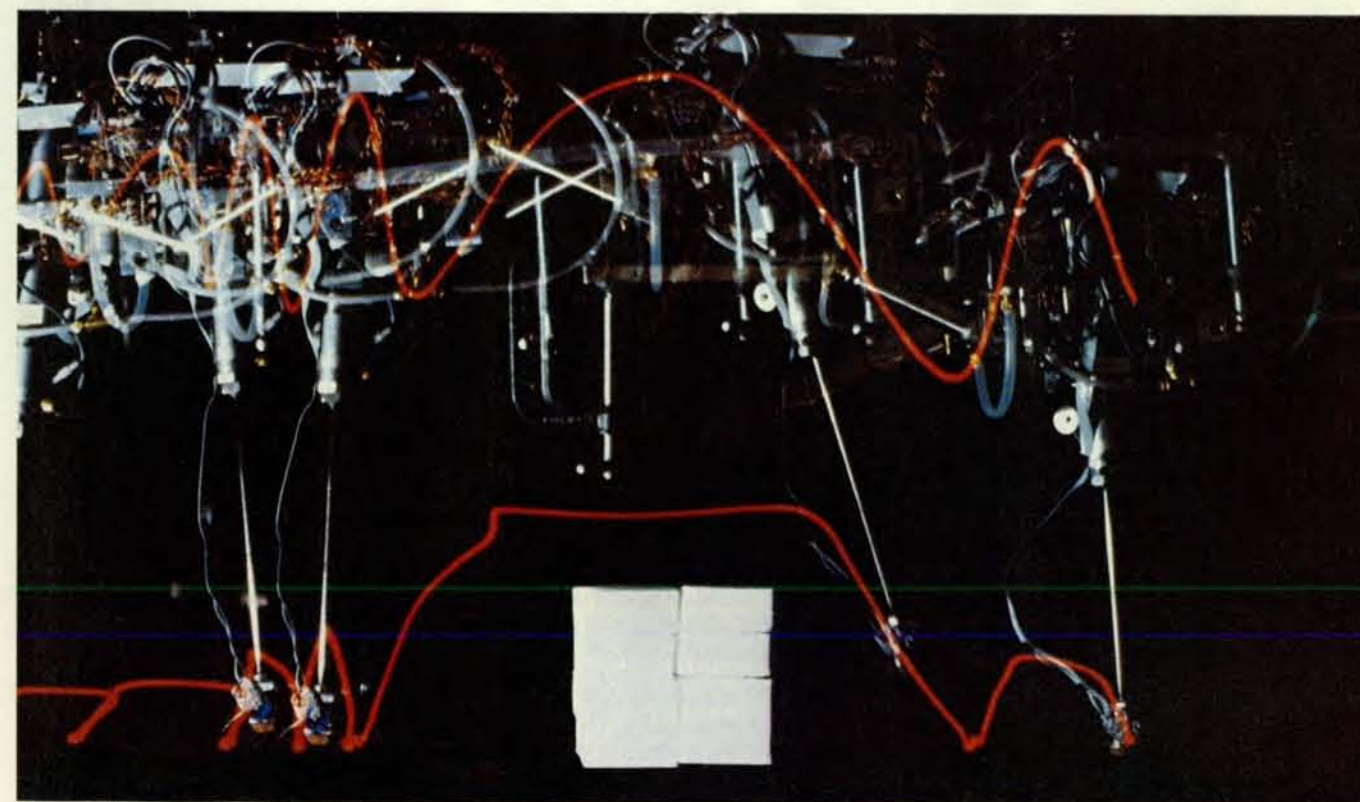
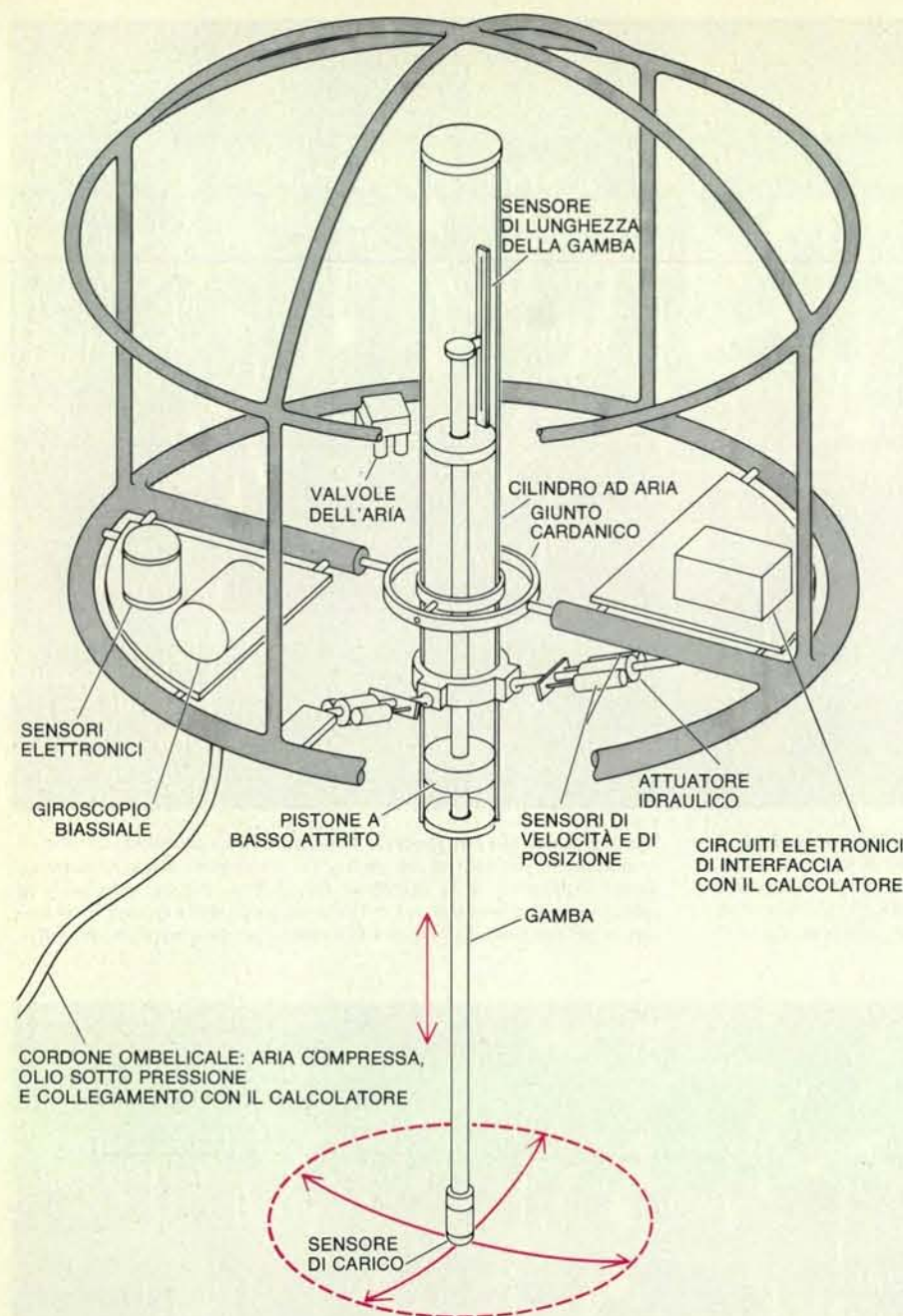
Le zampe di un animale durante la corsa oscillano fortemente in avanti e all'indietro per assicurare l'equilibrio e l'avanzamento. Abbiamo trovato che tali moti oscillatori non devono essere esplicitamente programmati, ma sono un esito naturale delle interazioni fra gli elementi di controllo dell'equilibrio e dell'assetto. Si supponga che il veicolo si sposti a velocità orizzontale costante e che atterri con il corpo in posizione verticale. Per mantenere la posizione eretta durante la fase di appoggio è necessario che il dispositivo di controllo dell'assetto impedisca la formazione di coppie di forze in corrispondenza dell'articolazione dell'anca. Poiché durante questa fase il piede è appoggiato al terreno, la gamba deve ruotare all'indietro di un certo angolo per far sì che la coppia di forze sull'articolazione sia nulla mentre il corpo si sposta in avanti.

D'altra parte come deve comportarsi il servomeccanismo per mantenere la fase di sospensione? Poiché il piede deve impiegare pressappoco lo stesso tempo per percorrere gli spazi davanti e dietro al centro di gravità del veicolo, la velocità di spostamento e la durata della fase di appoggio impongono al piede una posizione per l'atterraggio tale da porlo in un punto adatto per il successivo periodo di appoggio. Durante ogni fase di sollevamento pertanto la gamba deve spostarsi in avanti su comando del servomeccanismo di equilibrio, e in ogni fase di appoggio deve inclinarsi all'indietro su comando del servomeccanismo di assetto; le oscillazioni in avanti o all'indietro necessarie per la corsa si ottengono in modo automatico per l'influsso reciproco dei due circuiti di equilibrio e di assetto.

Stiamo ora costruendo una versione della macchina che potrà restare in equilibrio in tre dimensioni e che sarà quindi in grado di muoversi in uno spazio aperto. Abbiamo redatto e verificato un programma per simulare al calcolatore il movimento di una simile macchina e abbiamo trovato che il controllo tridimensionale può essere riportato agli stessi tre circuiti di retroazione già descritti.

Il nostro lavoro per la realizzazione della macchina a una sola gamba è stato molto facilitato da un'idea sorta durante la sua esecuzione: la corsa può essere meglio compresa suddividendola nelle tre parti esaminate in precedenza e cioè i controlli di altezza, di equilibrio e di assetto. La ripartizione del controllo fra queste tre funzioni ha reso assai più comprensibile il complesso comportamento delle gambe durante il cammino e la corsa. L'intuizione ha condotto a un sistema di controllo abbastanza semplice, che permette alla macchina di funzionare in equilibrio.

Il successo conseguito ci incoraggia a intraprendere la costruzione di macchine semoventi dotate di più gambe. Riteniamo che per affrontare correttamente il problema si debba concentrare l'attenzione anzitutto sull'equilibrio e sul moto verticale, per passare solo in seguito allo studio delle complicazioni introdotte dal movimento in avanti. Il concetto che l'attività principale sia il salto era naturale per la macchina a una sola gamba e conduceva a un'efficace linea di pensiero sul suo comportamento; lo stesso concetto sembra però meno naturale per la macchina con più gambe. E forse sembra meno naturale perché siamo abituati a vedere correre gli animali e desideriamo comprendere il loro comportamento subito nel suo complesso.

Diverse sono le possibili successioni di attività delle gambe utilizzabili da una macchina a quattro gambe che salta sul posto. Lo schema più semplice sarebbe il salto contemporaneo sulle quattro gambe. Non è difficile immaginare che gli stessi tre servomeccanismi che controllano la macchina che salta con una sola gamba potrebbero controllare l'analogo movimento effettuato da un veicolo a quattro gambe. In effetti il circuito del controllo di assetto che mantiene in posizione eretta il corpo sarebbe notevolmente più semplice grazie alla maggiore ampiezza della base di sostegno. Quando la macchina dovesse avanzare, le gambe oscillerebbero insieme secondo uno schema di movimento che potrebbe essere generato con facilità dagli stessi meccanismi di controllo che vengo-

Uno degli autori (Raibert) ha costruito una macchina che avanza a salti, per esaminare i problemi relativi al controllo di un veicolo che deve tenersi in equilibrio mentre si muove. L'unica gamba è azionata ad aria compressa e i movimenti sono controllati da un calcolatore al quale giungono segnali di retroazione da sensori di posizione. La versione illustrata qui è supportata da un braccio e l'equilibrio deve essere quindi tenuto solo in un piano. La fotografia, ottenuta tenendo aperto l'obiettivo della macchina fotografica, registra una serie di salti. Sono state montate sul corpo e sul piede della macchina alcune luci rosse che hanno permesso di evidenziare la traiettoria dei balzi.

La macchina a una sola gamba è in grado, con il suo movimento a balzi, di superare un ostacolo. Il veicolo si avvicina all'ostacolo da destra e quando dista un passo, l'operatore preme un apposito pulsante: nell'attuatore di comando si genera la tensione massima, per aumentare l'altezza del salto successivo. La gamba sollevata viene retratta e il suo moto normale di oscillazione è rallentato, per assicurare il superamento dell'ostacolo. Un servomeccanismo, che controlla l'equilibrio, sposta la gamba per farle assumere la corretta angolazione e la estende in previsione dell'atterraggio, dopo di che la macchina continua nel suo normale movimento a balzi. L'ostacolo era alto 15 centimetri.

Questa macchina per saltare avanzata, attualmente in fase di sviluppo, è stata progettata per operare in tre dimensioni. Essa è alta un metro circa, pesa 20 chilogrammi ed è collegata a un calcolatore situato nelle sue vicinanze. L'energia per il moto viene fornita da aria compressa, che regola inoltre l'altezza dei salti. Gli attuatori che determinano la posizione del piede sono idraulici e durante il balzo agiscono in modo da mantenere la macchina in equilibrio. Quando il piede appoggia al terreno gli attuatori mantengono il corpo della macchina in posizione eretta.

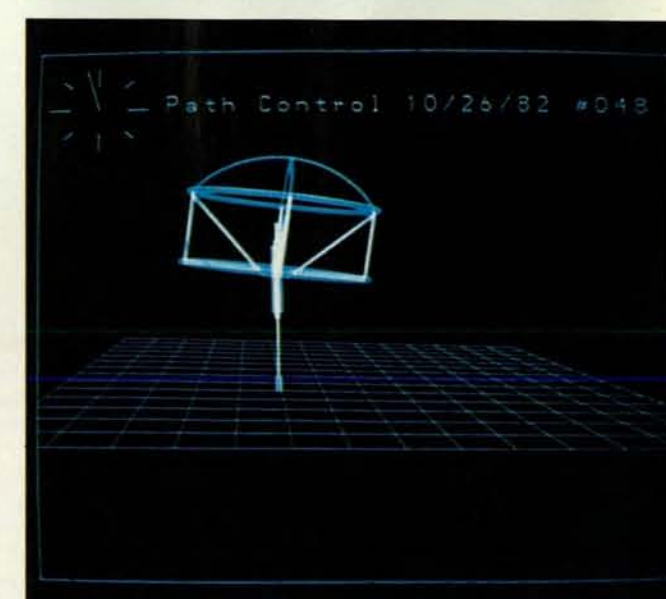
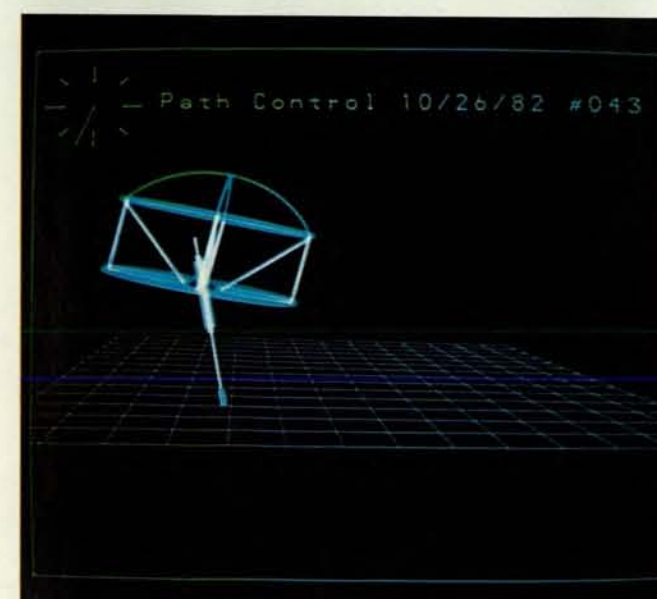
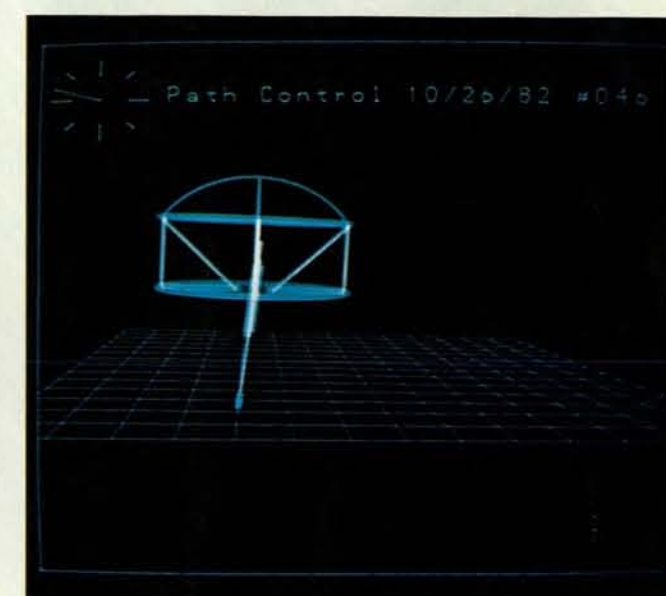
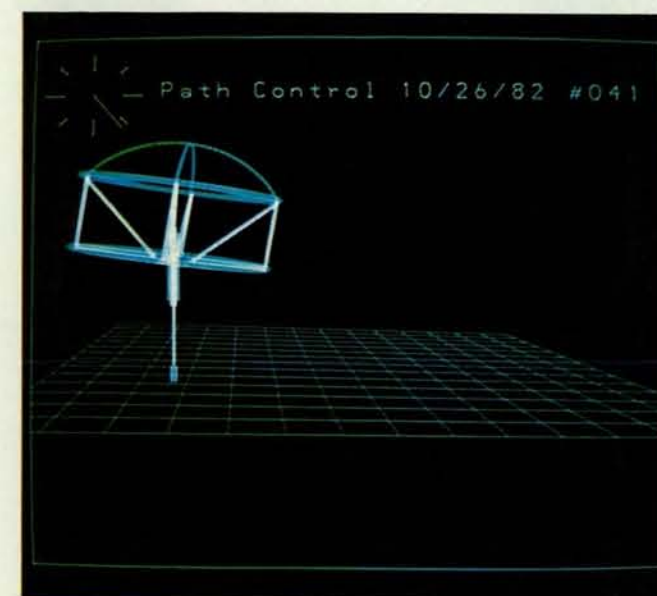
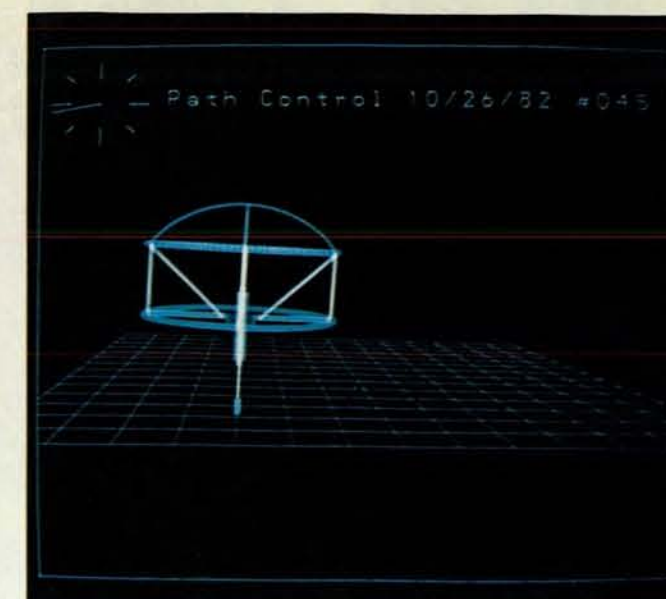
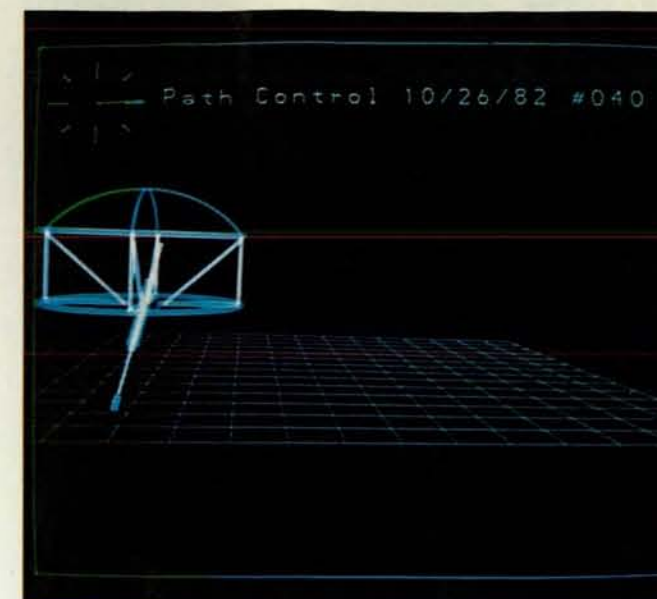
no impiegati nella macchina dotata solamente di una gamba.

Un altro possibile modo di saltare sul posto, per una macchina a quattro gambe, è di «rimbalzare» alternativamente sulle coppie di piedi diagonalmente opposti. Anche in questo caso non occorre troppa immaginazione per comprendere come sia possibile suddividere il controllo di un simile veicolo nei controlli di altezza, equilibrio e assetto. Il controllo di altezza dovrebbe erogare energia durante il movimento di salto per mantenere l'altezza dei salti al livello desiderato. Il controllo di equilibrio dovrebbe stabilire la posizione delle gambe alzate in modo da mantenere l'equilibrio. Il controllo di assetto del corpo dovrebbe applicare le corrette coppie di forze alle coppie di gambe che appoggiano al terreno. I controlli di equilibrio e di assetto alternerebbero l'uso degli stessi attuatori proprio come avviene nel caso della macchina a una sola gamba. Inoltre, come si è visto per quest'ultima, il moto di avanzamento potrebbe essere facilmente realizzato spostando in avanti ogni piede alzato verso una posizione scelta in maniera da rendere nulla la risultante delle forze che agiscono sul piede stesso durante la successiva fase di appoggio. Il movimento ottenuto sarebbe un trotto analogo a quello comune per molti animali a quattro zampe.

Altri due tipi di andatura possono essere compresi in modo analogo separando il controllo di ciascuna gamba nelle tre componenti: verticale, di equilibrio e di assetto. Durante il galoppo le zampe posteriori toccano terra poco prima di quelle anteriori e l'assetto del corpo viene modificato durante la fase sollevata in maniera tale che l'animale appoggia al terreno le zampe posteriori a muso alzato e alza le zampe anteriori a muso basso. La corsa a balzi è una variazione del galoppo nella quale le zampe anteriori operano pressoché all'unisono, come avviene per quelle posteriori, mentre i movimenti delle zampe anteriori e posteriori avvengono in uguali intervalli di tempo: è questo il tipo di andatura che consente al ghepardo di raggiungere velocità di oltre 95 chilometri all'ora.

Perché il movimento sul terreno avvenga con buon rendimento occorre che durante il moto la macchina perda poca energia. Si è già detto che si può aumentare il rendimento dei movimenti verticali delle gambe immagazzinandone l'energia cinetica in elementi elastici, ma come fare per i movimenti di oscillazione?

Quando la velocità rispetto al terreno è elevata, le gambe di una macchina devono muoversi avanti e indietro molto rapidamente. La maggior parte dell'energia spesa da un animale durante la corsa viene impiegata per generare tale moto delle gambe. Nella macchina a una sola gamba questi movimenti sono comandati da un servomeccanismo proporzionale, un sistema nel quale l'energia cinetica della gamba durante l'oscillazione viene del tutto perduta quando la gamba si arresta momentaneamente a fondo corsa. Al contrario il movimento di salto è realizzato con un sistema autorisonante costituito dalla molla della gamba e dalla massa della macchina, cosicché il servomeccanismo dell'altezza deve solo aggiungere o sottrarre una piccola quantità di energia per mantenere l'altezza del salto. È ovvio che, per poter avere un rendimento elevato ed essere nello stesso tempo veloce, una macchina a più gambe deve possedere qualche tipo di sistema autorisonante anche per il movimento oscillatorio delle gambe. Si potrebbe progettare un meccanismo comprendente delle molle poste fra le gambe per far oscillare queste ultime come un diapason su una frequenza adatta alla velocità verticale di salto.

Anche se riteniamo di aver compreso come costruire una macchina a quattro gambe in grado di muoversi secondo una qualsiasi delle andature descritte, rimangono aperti diversi problemi connessi con la partenza e l'arresto di una macchina del genere e con la scelta dell'andatura. È facile comprendere come avviare l'avanzamento della macchina quando questa salta già sul posto. Molto meno ovvio è il coordinamento della transizione fra una partenza da fermo e una corsa a velocità mas-

Le fotografie mostrano il movimento simulato della macchina a una sola gamba in tre dimensioni ripreso dallo schermo del calcolatore che ha predisposto i movimenti. Nella sequenza, che dura circa 0,7 secondi, la macchina si sta tenendo in equilibrio durante l'atterraggio e il salto successivo.

sima. Analogamente, come e quando una macchina del genere dovrebbe passare da un'andatura a un'altra? Un cavallo in corsa quando gira, cambia passo: se prima la zampa anteriore sinistra precedeva leggermente la destra, ora le posizioni sono invertite. Quali calcoli occorre eseguire per effettuare tali piccoli cambiamenti nello schema di movimento delle gambe? Ci sembra che questi problemi siano importanti, sia posti sotto forma di una domanda tecnica, «cosa costruire?» sia formulati come una domanda scientifica, «come si comportano i sistemi viventi?».

Un problema più difficile è la scelta dei punti di appoggio della macchina. Infatti ancora non sappiamo quale funzione abbiano, nell'uomo e negli animali, durante il cammino e la corsa, la vista e in particolare la capacità di scegliere i punti più adatti per appoggiare i piedi. Si potrebbe pensare di aggirare il problema facendo correre la macchina velocemente solo su terreni regolari e fornendole qualche assistenza umana per scegliere il percorso più sicuro. È probabile che si debba esplorare visivamente il suolo davanti al veicolo per individuare le asperità. In effetti come un cavallo al galoppo corre il rischio di infilare uno zoccolo nella tana di qualche animale, così la macchina corre un analogo rischio durante la corsa.

La mobilità dei veicoli da fuori-strada è limitata da due fattori. In primo luogo ruote e cingoli non consentono di sfruttare punti di appoggio discontinui, accessibili anche a un veicolo dotato di gambe. Siamo incoraggiati a continuare nel nostro lavoro dal pensiero che lo sviluppo dei calcolatori è oggi sufficientemente avanzato per permettere la costruzione di sistemi di controllo adeguati ai veicoli dotati di gambe e quindi per considerare con serietà questa alternativa nella realizzazione di veicoli ad alta mobilità. In effetti la Defense Advanced Research Projects Agency ha già finanziato ricerche su tali veicoli, e in parte anche i nostri progetti.

Un secondo fattore di mobilità è costituito dalla larghezza: una motocicletta può infilarsi in spazi ove una jeep non passa. È possibile costruire veicoli dotati di gambe e stretti; è però necessario che possano controllare l'equilibrio almeno nelle direzioni laterali. Essendo coinvolti sia nella costruzione di una macchina a sei gambe in grado di muoversi con una andatura radente senza problemi di equilibrio, sia in studi sul cammino e sulla corsa, ove l'equilibrio è essenziale, riteniamo che l'aspetto più importante della ricerca sia proprio l'equilibrio. Pensiamo che gli esperimenti sulle macchine a sei gambe ad andatura radente, in corso nel nostro laboratorio e altrove, siano soprattutto esercizi di controllo delle diverse gambe, ma che non siano di per se stessi di grande utilità; queste macchine prima o poi saranno sostituite da altre dotate di un minor numero di gambe, in grado di tenersi in equilibrio. La padronanza dell'equilibrio sarà la chiave per la costruzione di macchine ad alta mobilità capaci di camminare e di correre.

La fisica delle canne d'organo

La maestosità del suono di un organo a canne è dovuta all'interazione, accuratamente in fase, di un getto d'aria che fluisce attraverso la bocca di ogni canna e la colonna d'aria che risuona entro la canna

di Neville H. Fletcher e Suszanne Thwaites

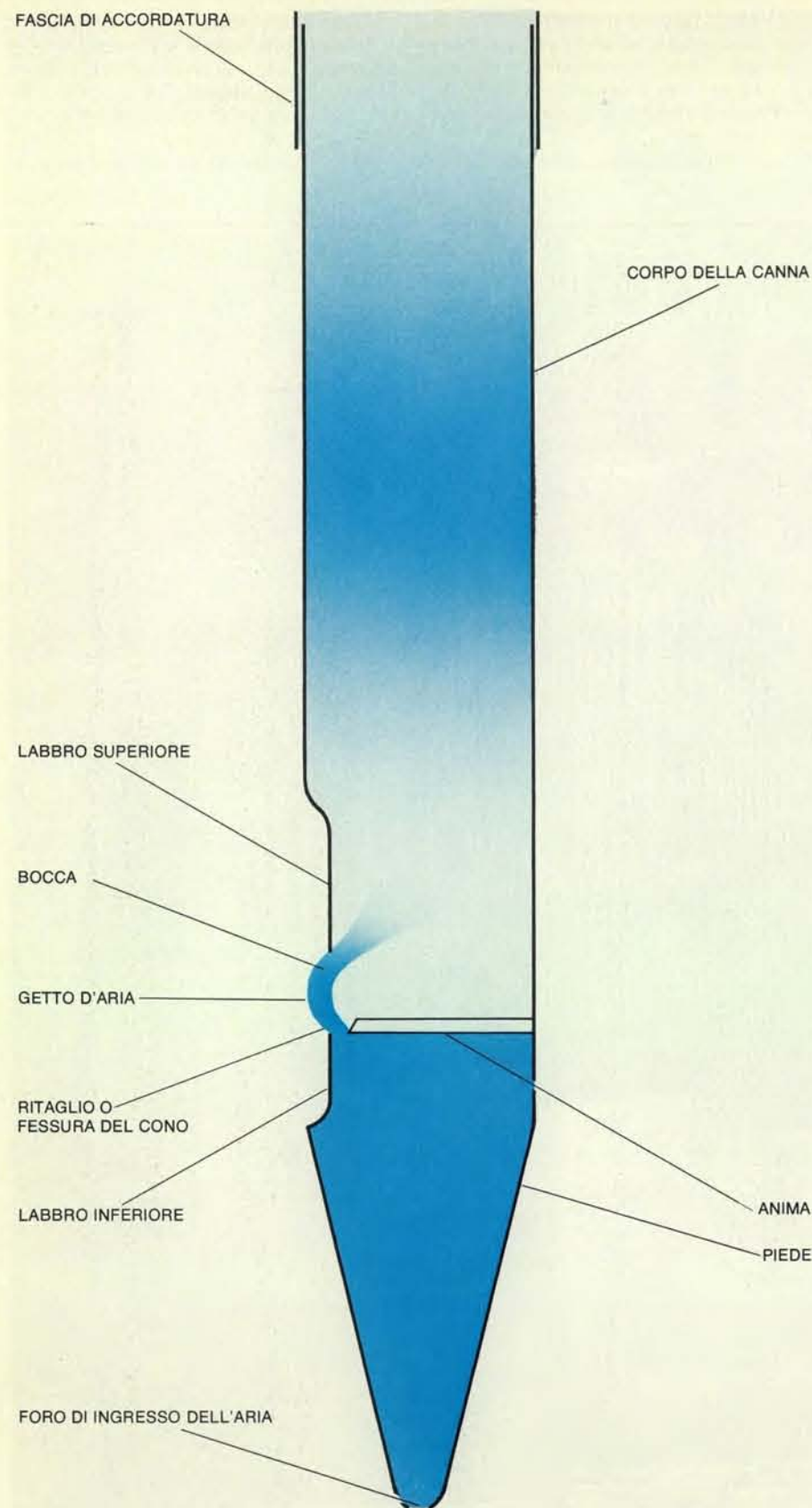
Nessun altro strumento musicale può reggere il confronto con l'organo a canne per potenza, timbro, dinamica, complessità e vera e propria maestosità di suono. Come molti altri strumenti musicali, l'organo è stato portato a un elevato grado di perfezione da generazioni di artigiani, attraverso il lento accumularsi della conoscenza empirica, fino a raggiungere, sostanzialmente, la sua forma moderna verso la fine del diciassettesimo secolo. Due dei più grandi fisici dell'Ottocento, Hermann von Helmholtz e Lord Rayleigh, hanno raggiunto conclusioni opposte in merito al meccanismo fondamentale che presiederebbe alla generazione del suono nell'organo a canne, ma sono mancati loro i mezzi tecnici per risolvere il problema. Grazie agli oscilloscopi e a molti altri dispositivi moderni, è stato possibile raggiungere una comprensione particolareggiata di quel meccanismo, e si è scoperto che le analisi di Helmholtz e di Rayleigh sono valide entrambe, ciascuna relativamente a determinati valori della pressione con cui l'aria viene immessa nelle canne. La spiegazione attuale del fenomeno, che esporremo in queste pagine, contrasta con molte spiegazioni del meccanismo di produzione del suono nelle canne d'organo, che possono ancora essere trovate in molti manuali.

I primi strumenti musicali furono probabilmente ottenuti da canne di palude o da altre piante dal fusto cavo: il suono poteva essere emesso soffiando trasversalmente all'estremità della canna, soffiando nell'estremità in modo da mettere in vibrazione le labbra, oppure stringendo l'estremità in modo da mettere in vibrazione le pareti della canna stessa, soffiando in essa. Flauti, trombe e clarinetti sono versioni moderne di questi tre primitivi strumenti a fiato, e tutti sono stati perfezionati in modo che un esecutore possa produrre un numero elevato di note diverse, coprendo una gamma notevole di frequenze acustiche.

Gli strumenti a canne hanno visto uno sviluppo parallelo, rappresentato dall'attribuzione di una canna diversa a ciascuna nota. Il più semplice degli strumenti di questo genere è il flauto di Pan, munito comunemente di circa 20 canne di lunghezza diversa, tutte tappate a una estremità e messe in vibrazione soffiando trasversalmente all'estremità aperta. All'altro estremo, per dimensioni e complessità, sta l'organo a canne, che può essere dotato anche di 10 000 canne, che l'esecutore controlla mediante complessi collegamenti meccanici. L'organo a canne ha una lunga storia. Sono giunte fino a noi figurine in ceramica di musici alessandrini che suonano una serie di canne messe in vibrazione mediante mantici, risalenti al secondo secolo a.C. Nel decimo secolo l'organo cominciò a entrare nell'uso liturgico delle chiese cristiane e per l'Europa cominciarono a circolare trattati di arte organaria scritti da monaci. Il grande organo costruito nel decimo secolo per la cattedrale di Winchester in Inghilterra si tramanda avesse 400 canne di metallo, 26 mantici e due tastiere per un totale di 40 tasti, ciascuno dei quali controllava 10 canne. Nei secoli successivi si ebbero progressi sostanziali dell'arte organaria, sia sotto il profilo meccanico sia sotto quello musicale e già nel 1429 fu costruito, per la cattedrale di Amiens, un organo con 2500 canne. Verso la fine del diciassettesimo secolo i costruttori tedeschi portarono lo strumento fino alla sua forma moderna.

La nomenclatura delle canne d'organo riflette le origini di questo strumento nei primitivi strumenti a fiato. Le canne d'organo sono aperte all'estremità superiore, mentre all'estremità inferiore si restringono a forma di cono, con una «bocca» (un'apertura) che corre orizzontalmente in una sezione appiattita immediatamente al di sopra del cono. All'interno della canna è alloggiata una «anima» (una lastrina orizzontale) con un «ritaglio» (una fessura sottile) tra l'anima stessa e il «labbro» inferiore della bocca. L'aria che mette in vibrazione la canna, fornita da un grande mantice, raggiunge il piede conico della canna con una pressione tipica di 500-1000 pascal. Inserendo il registro appropriato e schiacciando il tasto corrispondente, l'aria viene immessa nella canna, fluisce verso l'alto e forma un getto laminare emergendo dal ritaglio del cono. Il getto fluisce attraverso la bocca e va a colpire il labbro superiore, dove interagisce sia con il labbro stesso, sia con la colonna d'aria nel corpo della canna, per conservare l'oscillazione regolare che genera il suono della canna. Il problema di come la canna possa operare la improvvisa transizione dal silenzio a un suono regolare è tutt'altro che semplice ed è degno di attenzione per se stesso, ma qui non lo prenderemo in esame. Ci occuperemo principalmente dei processi che controllano il suono regolare delle canne d'organo e determinano le loro caratteristiche qualità timbriche.

Il comportamento del getto che dà inizio al suono e lo alimenta sembrerebbe presentare un problema diretto di fluidodinamica. Si scopre, tuttavia, che anche un getto regolare che fluisce con continuità in modo laminare risulta molto difficile da trattare sotto il profilo teorico, mentre il getto completamente turbolento che fluisce in realtà in una canna d'organo è di una complessità incredibile. Per fortuna, però, la complessità introdotta dalla turbolenza in realtà semplifica il comportamento del getto. Se il flusso fosse laminare, l'interazione fra il getto e l'ambiente circostante sarebbe governata da forze viscose. Nel getto reale, invece, il meccanismo di interazione non è la viscosità ma la turbolenza, e su una scala in rapporto diretto con la larghezza del getto stesso. I costruttori di organi, per essere sicuri che i getti d'aria nelle canne dei loro strumenti siano completamente turbolenti, ricavano dei sottili denti lungo il bordo dell'anima. Paradossalmente, un getto turbolento è stabile e riproducibile, mentre un getto laminare non lo è.

L'organo della sala concerti della Sydney Opera House, completato nel 1979, è uno dei più grandi del mondo e dei più perfezionati. Progettato e costruito da Ronald Sharp, possiede circa 10 500 canne distribuite in 205 file e controllate dall'azione meccanica di cinque tastiere e una pedaliera. L'azione meccanica, che regola il flusso di aria nelle canne, è duplicata da un'azione elettrica controllata da microelaboratori. L'organo quindi può essere azionato da un nastro magnetico su cui sia stata registrata in forma digitale una esecuzione originale riproducendola.

Una canna d'organo è eccitata dall'aria che vi viene immessa dal basso. L'aria assume la forma di un getto attraversando il ritaglio del cono, tra il labbro inferiore e l'anima (una lastrina orizzontale). Quando l'aria attraversa la bocca della canna, interagisce con la colonna d'aria che già vi si trova all'altezza del labbro superiore, fluendo alternativamente dentro e fuori la canna stessa. Le onde che si propagano lungo il getto turbolento mantengono una oscillazione regolare nella colonna d'aria, determinando il suono. Lo schema a onda stazionaria della pressione nella canna è rappresentato dal chiaroscuro in colore. L'estremità superiore della canna è circondata da una fascia regolabile, come in questo disegno, oppure è chiusa da un cocuzzolo regolabile: ambedue i dispositivi consentono una lieve modifica della lunghezza della colonna d'aria, per l'accordatura.

Un getto completamente turbolento si mescola gradualmente con l'aria circostante, e di conseguenza si allarga e va rallentando in modo semplice. Quando si rappresenta in grafico la velocità del getto in funzione della distanza dal suo piano centrale, la curva risultante è a forma di campana, con la velocità massima al centro. La larghezza del getto aumenta linearmente con la distanza dal ritaglio del cono. La quantità di moto del flusso si deve conservare, il che significa che la velocità diminuisce con la radice quadrata della distanza dal ritaglio. Questa descrizione può essere sostenuta con motivazioni teoriche e si accorda con i dati sperimentali (quando si include una piccola regione di transizione in prossimità del ritaglio del cono).

In un organo a canne che già emette un suono il getto emerge dal ritaglio del cono per entrare nell'intenso campo sonoro della bocca della canna. Il movimento dell'aria associato con il suono è diretto dentro e fuori della bocca e pertanto in direzioni perpendicolari al piano del getto. Mezzo secolo fa Burniston Brown dell'University College di Londra riprese splendide fotografie di getti laminari resi visibili con fumo, che emergevano in un campo sonoro, e scoprì che sviluppavano onde sinuose, le quali crescevano percorrendo il getto fino a che il getto non si rompeva in una doppia fila di anelli vorticosi rotanti in direzioni opposte. L'applicazione un po' ingenua di queste e di altre osservazioni consimili è responsabile di molta della confusione che si incontra tuttora, a proposito della fisica delle canne d'organo, in molti manuali.

Un metodo più fruttuoso per studiare il comportamento dei getti reali in un campo sonoro è quello di togliere la canna d'organo e di generare il campo sonoro per mezzo di un altoparlante. Studi di questo tipo, condotti da John W. Coltman della Westinghouse Electric Corporation e dal nostro gruppo all'Università del New England in Australia, hanno dato un importantissimo contributo alla nostra conoscenza attuale della fisica delle canne d'organo. In realtà Rayleigh aveva sviluppato una descrizione matematica precisa e pressoché completa del comportamento dei getti laminari nei fluidi non viscosi. Poiché fortunatamente, come abbiamo visto, la turbolenza semplifica il comportamento del getto, anziché complicarlo, si può estendere la trattazione di Rayleigh, con poche modificazioni, in modo da descrivere i getti reali prodotti e analizzati sperimentalmente da Coltman e da noi.

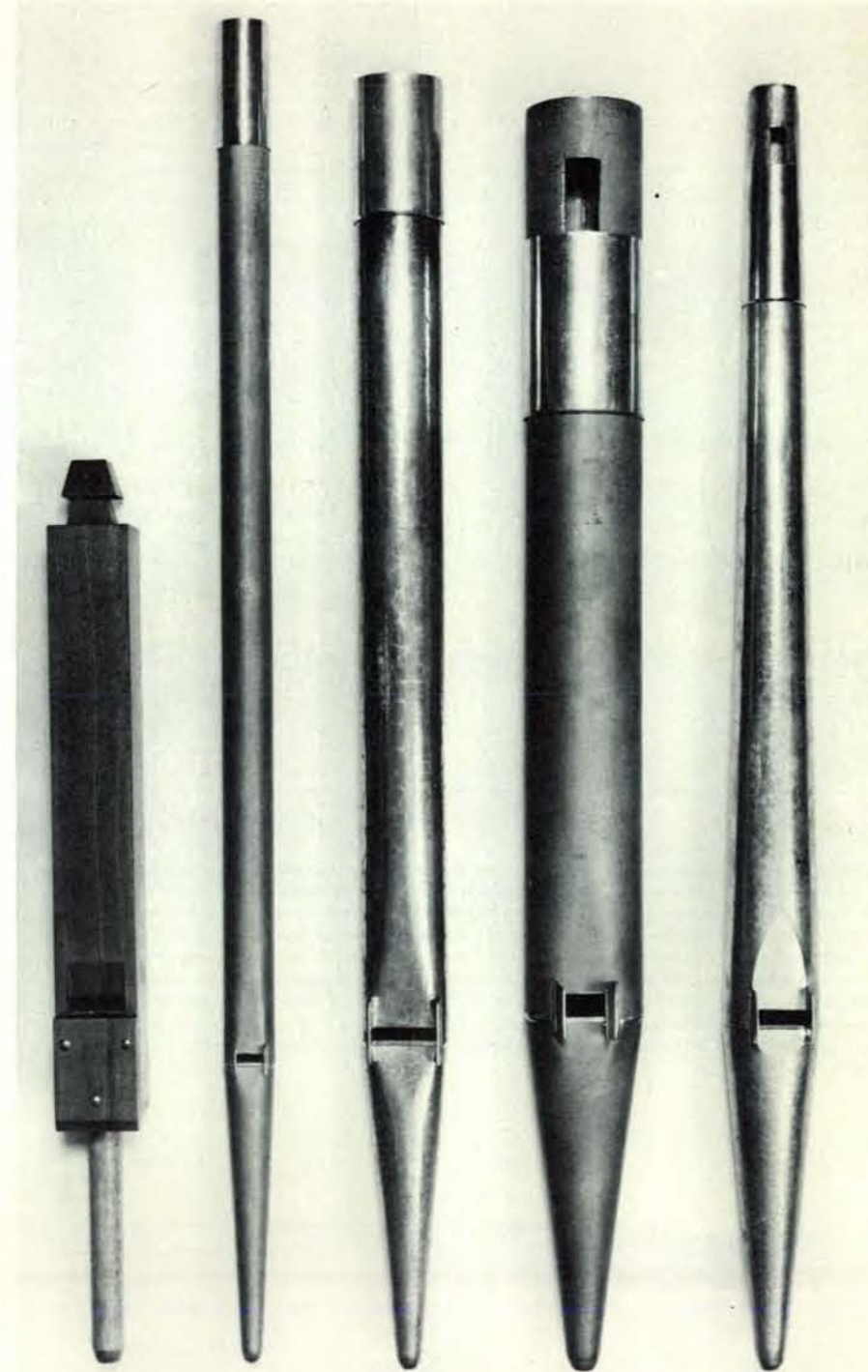
Se ci si scordasse del ritaglio del cono nella canna d'organo, ci si aspetterebbe che il getto, essendo semplicemente uno strato di aria in movimento, si muova semplicemente avanti e indietro, per le vibrazioni acustiche, insieme a tutta l'altra aria che si trova nella bocca della canna. Quando il getto lascia il ritaglio, tuttavia, viene efficacemente mantenuto calmo dal ritaglio stesso; al movimento generale di oscillazione avanti-indietro si sovrappone uno spostamento localizzato al ritaglio, che genera un perfetto equilibrio. Lo spostamento localizzato, che corrisponde perfettamente, per frequenza e ampiezza, al campo sonoro, per conservare uno spostamento nullo al ritaglio, viene trasportato con l'aria in moto nel getto e impone a questo un movimento ondulatorio sinuoso.

Come Rayleigh ha dimostrato per un caso limite nei suoi getti, e come noi abbiamo verificato nei particolari per getti turbolenti divergenti, il moto ondulatorio si propaga lungo il getto a una velocità di poco inferiore alla metà della velocità dell'aria nel piano centrale del getto. Va aggiunto, tuttavia, che l'onda cresce in ampiezza quasi esponenzialmente, mentre si propaga nel getto. In un caso tipico l'ampiezza dell'onda raddoppia mentre copre circa un millimetro del suo percorso nel getto, cosicché rapidamente gli effetti dell'onda dominano il semplice movimento laterale di avanti-indietro imposto dalle vibrazioni acustiche.

Troviamo che il tasso di crescita dell'onda è massimo quando la sua lunghezza d'onda nel getto è circa sei volte la larghezza del getto stesso in quel punto. Invece, quando la lunghezza d'onda è inferiore alla larghezza del getto, l'onda non cresce affatto e può addirittura morire. Poiché il getto rallenta e si allarga mentre si allontana dal ritaglio del cono, solo le onde lunghe, cioè quelle di bassa frequenza, possono propagarsi con una ampiezza elevata nei getti lunghi. L'importanza di questo fatto risulterà evidente quando prenderemo in considerazione il contenuto armonico dei suoni delle canne d'organo.

Consideriamo ora gli effetti del campo sonoro della canna d'organo sul getto. Come si può facilmente immaginare, le grandi onde acustiche associate con il campo sonoro nella bocca delle canne faranno sì che la punta del getto ondeggi avanti e indietro lungo il labbro superiore della bocca, e di conseguenza il getto fluirà in alternanza dentro e fuori la canna. La situazione non è dissimile da quella che si verifica quando si spinge un'altalena che già si muove. La colonna d'aria nella canna sta già oscillando e se gli sbuffi d'aria entrano nella canna in fase con le oscillazioni, essi agiranno nel senso di mantenere le oscillazioni contro le numerose perdite di energia, dovute a cause come l'irraggiamento del suono della canna e le forze di resistenza contro le pareti della canna. Se, invece, gli sbuffi non sono in fase con le oscillazioni della colonna d'aria, il getto tenderà a smorzare le oscillazioni e il suono si affievolirà fino a morire.

La relazione fra i movimenti acustici dell'aria nella bocca della canna e il momento in cui l'impulso d'aria arriva all'interno del labbro superiore sono determinate semplicemente dal tempo necessario perché un'onda nel getto viaggi dal ritaglio del cono fino al labbro superiore. Gli organari chiamano lume questa distanza. Se il lume è grande o la pressione a cui viene insufflato il getto è bassa (e di conseguenza il getto ha bassa velocità), il tempo necessario all'onda per viaggiare dal ritaglio al labbro superiore sarà lungo. Al contrario, se il lume è piccolo o la pressione è elevata, il tempo di percorrenza sarà breve.

Al fine di scoprire quale sia esattamente la corretta relazione di fase fra l'oscillazione della colonna d'aria nella canna e il sopraggiungere degli impulsi del getto all'interno del labbro superiore, è necessario sapere qualcosa di più sul

Questa fotografia mostra cinque canne di forma diversa che danno tutte una nota della stessa altezza, ma con caratteristiche timbriche diverse. La seconda da sinistra è la dulciana, con un suono dolce e sottile che assomiglia a quello di uno strumento a corda. La terza è un principale aperto, dal suono brillante e pieno tipico dell'organo. La quarta ha il suono di un flauto molto dolce. La quinta è un *waldflöte* (letteralmente, flauto del bosco), dal suono morbido. La canna di legno a sinistra è chiusa da un cocuzzolo: ha la stessa frequenza fondamentale delle altre quattro canne, ma produce solo armoniche dispari, cioè sopratoni con frequenze multiple della fondamentale per un fattore dispari. Le altre canne differiscono leggermente l'una dall'altra in lunghezza, a causa delle piccole correzioni finali che sono state apportate per intonarle perfettamente.

modo in cui gli impulsi agiscono sulla colonna d'aria. Helmholtz pensava che il fattore dominante, in questa relazione, fosse il volume del flusso determinato dal getto. Allora, se gli impulsi nel flusso del getto dovessero trasmettere tutta l'energia possibile all'oscillazione della colonna d'aria, dovrebbero essere immessi nella canna allorché la pressione acustica all'interno del labbro superiore è massima.

Rayleigh seguì un'idea diversa. Poiché la bocca - questo il suo ragionamento - non è molto lontana dall'estremità aperta della canna, immediatamente al suo interno potrebbe generarsi una minima pressione acustica sufficiente a ostacolare il flusso del getto. Stimò che il flusso del getto venisse sostanzialmente bloccato non appena entra nella canna, generando quindi rapidamente una pressione che può agire sul flusso nella canna. Di conseguenza per Rayleigh il massimo trasferimento di energia dal getto si avrebbe se il getto entrasse nella canna quando è massimo il flusso acustico, e non quando è massima la pressione. La differenza fra questi due massimi è pari a un quarto del periodo di oscillazione nella frequenza di oscillazione della colonna d'aria nella canna. Per riprendere il paragone dell'altalena, la differenza è la stessa che corre fra il dare una spinta all'altalena quando si trova nella posizione più elevata e ha acquisito la massima energia potenziale (idea di Helmholtz) e il dare una spinta all'altalena quando si trova nella posizione più bassa ed è al massimo della sua velocità (idea di Rayleigh).

Il problema è rimasto senza soluzione (in effetti, nessuno se ne è più occupato) per quasi 80 anni. Oggi abbiamo una risposta abbastanza soddisfacente, grazie agli studi di Lothar Cremer dell'Istituto Heinrich Hertz per le ricerche sui fenomeni vibratori di Berlino, di Samuel A. Elder della United States Naval Academy, di Coltman e del nostro gruppo. Per dirla in breve, sia Helmholtz sia Rayleigh avevano parzialmente ragione. L'equilibrio tra i due meccanismi di spinta dipende dalla pressione a cui è insufflata l'aria e dalla frequenza del suono: il meccanismo di Helmholtz prevale a basse pressioni e frequenze elevate, mentre il meccanismo di Rayleigh prevale a pressioni elevate e basse frequenze. Per una canna d'organo normale, di solito il meccanismo di Helmholtz è più importante.

Coltman ha ideato un metodo semplice ed efficace per studiare le proprietà del getto; il suo metodo, poi, è stato esteso e modificato nel nostro laboratorio. In sostanza si esamina il comportamento di un getto alla bocca di una canna in cui siano stati inseriti cunei di feltro o di spugna all'estremità distale, in modo che la canna non emetta suono. Poi si dirige un'onda sonora nella canna mediante un altoparlante posto all'esterno dell'estremità distale. L'onda viene riflessa dall'estremità della bocca, con e senza getto. In ambedue i casi le onde sonore incidenti e riflesse interagiscono all'interno della canna creando uno schema di onde stazionarie. Misurando con un piccolo microfono sonda i cambiamenti dello schema quando viene attivato il getto, siamo in grado di stabilire se il getto fornisce energia all'onda riflessa, o ne assorbe da questa.

Quel che i nostri esperimenti determinano effettivamente è l'«ammettenza» del getto, definita come il rapporto fra il flusso acustico che esce dalla bocca, prodotto dalla presenza del getto, e la pressione acustica immediatamente all'interno della bocca. L'ammettenza acustica possiede sia valore assoluto sia un angolo di fase, che può essere espresso come funzione della frequenza o della pressione a cui viene insufflato il getto. Se si fanno variare separatamente o la frequenza o la pressione, riportando opportunamente in grafico l'ammettenza, si ottiene una forma a spirale *(si veda l'illustrazione in alto nella pagina a fronte)*. La distanza dall'origine della spirale rappresenta il valore assoluto dell'ammettenza, mentre la posizione angolare lungo la spirale rappresenta il ritardo di fase dell'onda sovraimposta al getto dalle oscillazioni acustiche nella canna. Un ritardo di una lunghezza d'onda corrisponde a 360 gradi lungo la spirale. Date le peculiari proprietà dei getti turbolenti, si ha che, se si moltiplica l'ammettenza per la radice quadrata della pressione a cui è insufflato il getto, tutte le misurazioni per una particolare canna d'organo vanno a cadere nella stessa spirale.

Se teniamo costante la pressione e aumentiamo la frequenza dell'onda ancora in ingresso, il punto che rappresenta l'ammettenza si muove a spirale verso il centro, in direzione oraria. Se manteniamo costante la frequenza e aumentiamo la pressione, il punto si muove a spirale verso l'esterno in verso antiorario.

Quando il punto che rappresenta l'ammettenza giace nella metà destra della spirale, il getto assume energia dal flusso della canna e di conseguenza si dissipa energia. Quando il punto giace nella metà di sinistra, il getto fornisce energia al flusso della canna: questo significa che il getto agisce come un generatore acustico. Quando il punto si trova nella metà superiore della spirale, il getto abbassa la frequenza naturale di risonanza della canna; quando il punto si trova nella metà inferiore, il getto alza la frequenza naturale di risonanza. L'angolo di riferimento dal quale si misura il ritardo di fase dipende dal meccanismo prevalente e, come abbiamo visto, il prevalere del meccanismo di Helmholtz o di quello di Rayleigh dipende dalla pressione e dalla frequenza. Tuttavia l'angolo non è mai molto lontano dalla posizione zero (le «tre» dell'orologio) sul semiasse orizzontale destro.

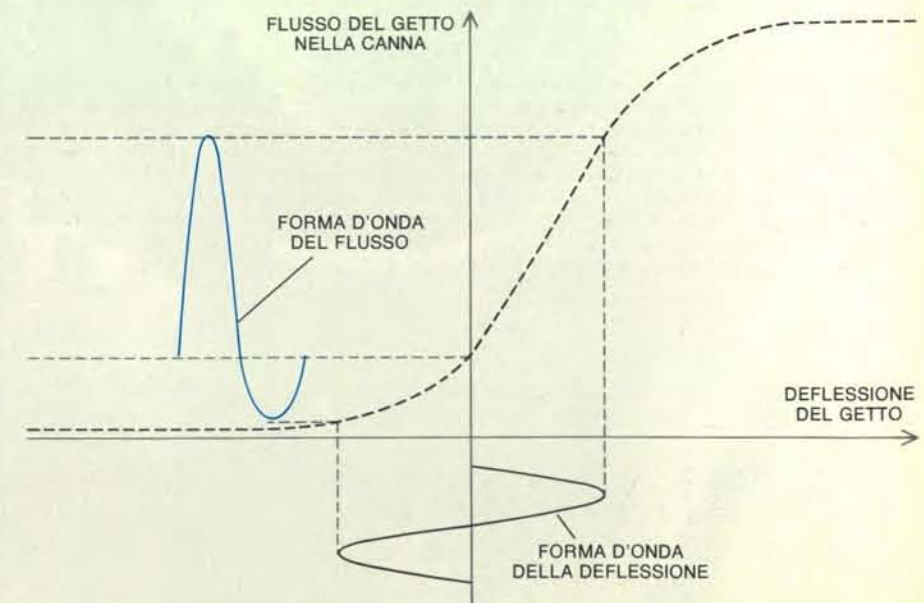
Poiché 360 gradi lungo la spirale rappresentano un ritardo di fase di una lunghezza d'onda nell'onda sinuosa che si propaga lungo il getto, si ha che per ritardi di fase che vanno da un po' meno di un quarto di lunghezza d'onda a quasi tre quarti di lunghezza d'onda un punto rappresentativo sulla spirale giacerà alla sinistra della linea centrale, cioè nella regione in cui il getto si comporta come un generatore acustico. Abbiamo anche visto che a frequenza costante il ritardo di fase è funzione della pressione, che controlla tanto la velocità del getto quanto la velocità di propagazione dell'onda sinuosa nel getto. Dal momento che la velocità dell'onda è pari alla metà della velocità del getto e la velocità del getto è proporzionale alla radice quadrata della pressione, la pressione deve modificarsi notevolmente perché la fase dell'onda nel getto vari di una mezza lunghezza d'onda. In effetti esiste un intervallo, pari a nove volte la pressione, all'interno del quale la canna può emettere il suo suono fondamentale, se tutte le altre condizioni sono favorevoli. In pratica, tuttavia, la canna passerà a un modo proprio di risonanza di frequenza superiore, prima che venga raggiunto il limite superiore di questo intervallo di pressione.

Va sottolineato che la spirale può avere più di un giro che si estende abbastanza a sinistra da superare le perdite dissipative nella canna e generare un suono stabile. Il secondo di questi giri, che in genere è l'unico, oltre al primo, che possa determinare l'emissione di suono da parte della canna, corrisponde a circa tre lunghezze d'onda e mezza nel getto. Poiché il getto ha solo una piccola ammettenza, a questo punto il suono generato è debole rispetto al suono che si genera per un punto nel giro più esterno della spirale.

La spirale dell'ammettenza di cui stiamo parlando presenta una complicazione ulteriore, se la deflessione subita dal getto in corrispondenza del labbro superiore diventa maggiore della larghezza del get-

Una serie di «istantanee» in successione mostra la forma del getto d'aria mentre emerge dal ritaglio del cono nel campo acustico generato nella bocca della canna dalla colonna d'aria risonante che si trova nel corpo della canna stessa. Lo spostamento periodico dell'aria nella bocca crea un'onda sinuosa che si propaga con velocità pari alla metà della velocità dell'aria nel piano centrale del getto e cresce esponenzialmente fino al momento in cui la sua ampiezza non arriva a superare la larghezza del getto. Le divisioni orizzontali visualizzano le distanze percorse dall'onda del getto in quarti successivi del periodo di oscillazione *T*. Le linee si avvicinano quando il getto perde velocità. In una canna d'organo il labbro superiore si trova all'incirca nel punto in cui nel disegno figura la freccia, cosicché il getto fluisce alternativamente entro e fuori la canna.

Le proprietà di generazione acustica del getto possono essere misurate inserendo dei cunei di feltro o di spugna nell'estremità aperta della canna, per evitare che produca un suono, e introducendo un'onda sonora di piccola ampiezza mediante un altoparlante. L'onda sonora viene riflessa dall'estremità opposta della canna e interagisce alla bocca con il getto d'aria. L'interazione del getto con lo schema a onda stazionaria nella canna viene esaminata con un microfono mobile. In questo modo è possibile stabilire se il getto fornisce energia all'onda riflessa alla base della canna, o se invece assorbe energia da essa. Perché la canna possa produrre un suono, il getto deve cedere energia. Le misurazioni sono in termini di «ammettenza» acustica, cioè del rapporto fra flusso acustico che esce dalla bocca e pressione acustica immediatamente all'interno della bocca stessa. Riportando in grafico l'ammettenza in funzione di varie combinazioni di pressione dell'aria e di frequenza, si ottiene la curva a spirale raffigurata nel disegno in alto della pagina a fronte.

L'ammettenza acustica del getto, in funzione della frequenza o della pressione a cui è insufflata l'aria, assume la forma di una spirale in cui la distanza dall'origine rappresenta il valore assoluto dell'ammettenza, mentre la posizione angolare rappresenta la relazione di fase fra il flusso acustico in uscita dalla bocca della canna e la pressione immediatamente all'interno della bocca. Quando il flusso in uscita è in concordanza di fase con la pressione, l'ammettenza si trova nella metà di destra della spirale e il getto dissipa energia. Perché il getto possa agire come un generatore acustico, l'ammettenza deve trovarsi nella metà di sinistra della spirale: per questo il movimento avanti-indietro del getto deve essere spostato, ovvero in ritardo di fase rispetto alla pressione all'interno della bocca della canna. L'onda riflessa dal getto, allora, è maggiore dell'onda incidente. L'angolo di riferimento rispetto al quale si misura il ritardo di fase dipende da quale è il meccanismo di eccitazione dominante, il meccanismo proposto nel secolo scorso da Hermann von Helmholtz oppure quello proposto da Lord Rayleigh. Quando l'ammettenza cade nella metà superiore della spirale, il getto abbassa la frequenza naturale di risonanza della canna; quando invece cade nella metà inferiore della spirale, il getto innalza la frequenza di risonanza. Le dimensioni delle parti esterne della spirale sono state ridotte, nel disegno, per motivi di maggior chiarezza.

La curva che dà il flusso del getto in una canna *(curva tratteggiata)* per una deflessione data è asimmetrica rispetto al punto di deflessione nulla, perché il labbro della canna è studiato in modo da tagliare il getto in posizione leggermente spostata rispetto al suo piano centrale. Quando il getto viene deflesso in una forma sinusoidale semplice con un'ampiezza elevata *(curva continua in nero)*, il flusso del getto nella canna *(curva in colore)* si «satura» prima a una estremità del suo intervallo, dove il getto fluisce completamente al di fuori del labbro della canna. Per un'ampiezza ancora maggiore il flusso del getto si saturerà anche all'altra estremità, dove il getto entra completamente nella canna. Lo spostamento del labbro dà all'onda del flusso una forma asimmetrica con componenti di frequenza a tutti i multipli interi della frequenza di deflessione. A queste componenti armoniche si deve la pienezza di suono di una canna d'organo ben regolata.

to. Il getto allora fluisce quasi completamente dentro e fuori la bocca della canna a ogni ciclo e l'impulso che impartisce all'onda riflessa nella canna diventa indipendente da qualunque ulteriore incremento di ampiezza. L'efficienza del getto come generatore diminuisce proporzionalmente. La spirale dell'ammettenza semplicemente assume dimensioni minori, mentre l'ampiezza della deflessione del getto aumenta.

La perdita di efficienza del getto all'aumentare dell'ampiezza della deflessione viene accompagnata da un aumento delle perdite di energia nella canna. L'oscillazione della canna rapidamente si stabilizza a una ampiezza per cui l'energia fornita dal getto controbilancia esattamente l'energia persa nella canna. Cosa abbastanza sorprendente, si scopre che nella maggior parte dei casi le perdite di energia dovute alla turbolenza e alla viscosità sono molto maggiori di quelle dovute all'irraggiamento del suono dalla bocca e dall'estremità aperta della canna.

Questa immagine dell'interno dell'organo della Sydney Opera House mostra alcune delle 205 file di canne, in prevalenza di metallo e solo in parte di legno. La lunghezza della parte attiva delle canne raddoppia ogni dodici canne; il diametro raddoppia circa ogni sedici canne. I maestri costruttori di organi hanno stabilito, grazie a una lunga esperienza, le proporzioni per ottenere un equilibrio delle caratteristiche timbriche.

Il suono di una canna d'organo reale non si limita, ovviamente, a un'unica frequenza, ma possiede numerose componenti di frequenza superiore. Si può dimostrare che queste componenti sono tutte armoniche esatte della frequenza fondamentale, cioè multipli interi di tale frequenza. In condizioni regolari di insufflazione dell'aria, la forma dell'onda sonora, vista attraverso un oscilloscopio, rimane esattamente la stessa: anche la minima deviazione rispetto ai multipli interi esatti, per le frequenze delle componenti, determinerebbe un visibile cambiamento nella forma d'onda, per quanto lento.

Questo comportamento è interessante, perché le risonanze proprie della colonna d'aria in una canna d'organo, e più in generale per qualunque canna aperta, hanno frequenze che si discostano leggermente dagli esatti multipli armonici. Questo è dovuto al fatto che la lunghezza efficace della canna diminuisce leggermente all'aumentare della frequenza, causa il modificarsi del flusso acustico alle estremità aperte. Come vedremo, le armoniche del suono di una canna d'organo sono generate dalle interazioni fra il getto e il labbro della canna, così che la canna agisce in gran parte come un risonatore passivo, per quel che riguarda le armoniche superiori.

Le risonanze di una canna si sviluppano quando è massimo il movimento dell'aria alle estremità aperte della canna stessa; il che equivale a dire che l'ammettenza di una canna d'organo deve essere massima alla bocca. Se ne trae la conclusione che le risonanze di una canna aperta alla sua estremità distale si verificano a frequenze per le quali la lunghezza della canna è pari a un numero intero di mezze lunghezze d'onda del suono nell'aria. Se la frequenza fondamentale è f_1, le risonanze superiori sono $2f_1$, $3f_1$ e via dicendo. (In realtà tutte le frequenze delle risonanze superiori superano un poco questi valori ideali, come abbiamo notato poco sopra.)

Per una canna tappata, cioè chiusa, alla sua estremità distale, le risonanze si presentano quando la lunghezza della canna è pari a un numero dispari di quarti di lunghezza d'onda. Pertanto una canna tappata deve essere lunga solo la metà di una canna aperta, per produrre la stessa nota: le sue risonanze sono a f_1, $3f_1$, $5f_1$, e via di seguito.

Per tornare al getto della canna d'organo, abbiamo visto che i disturbi ad alta frequenza tendono a morire nel getto quando la larghezza del getto aumenta, così che la punta del getto al labbro superiore oscilla avanti e indietro quasi sinusoidalmente alla frequenza della componente fondamentale del suono della canna, in modo quasi del tutto indipendente dalle componenti di frequenza superiore nel campo acustico alla bocca. Un flusso sinusoidale del getto, tuttavia, non porta a un flusso sinusoidale del getto nella canna, poiché il flusso «si satura» fluendo completamente all'interno o all'esterno del labbro superiore a ogni estremità dell'intervallo di deflessione del getto. Oltre a questo, di solito il labbro è leggermente spostato, cosicché non taglia il getto esattamente lungo il suo piano centrale, il che rende asimmetrica la saturazione. Pertanto la forma d'onda del flusso del getto nella canna ha tutte le armoniche della frequenza fondamentale bloccate in una precisa relazione di frequenza e di fase, e le ampiezze relative delle armoniche di frequenza elevata aumentano rapidamente all'aumentare dell'ampiezza della deflessione del getto.

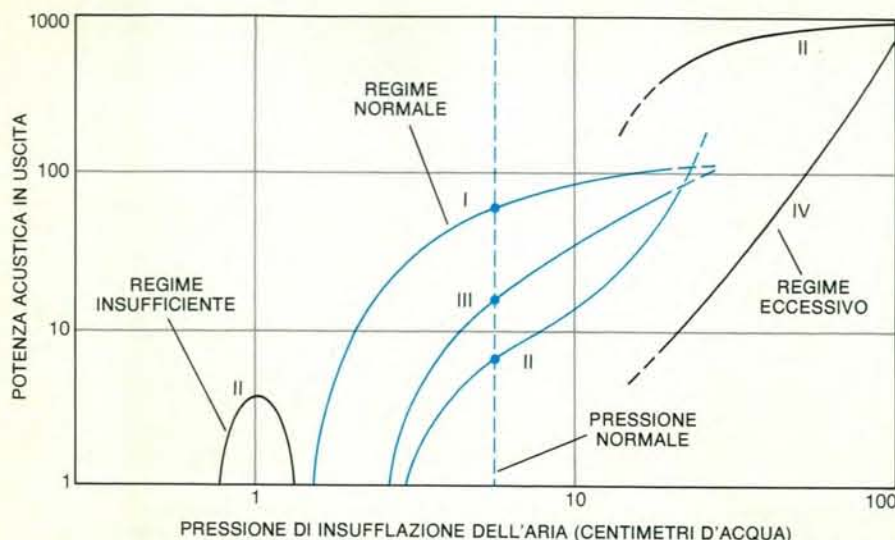
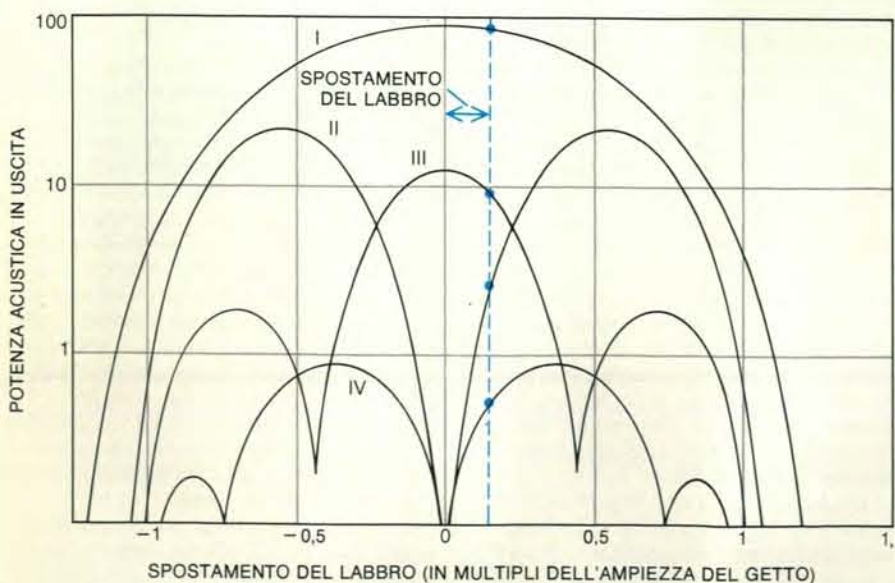
In una tipica canna d'organo la deflessione che il getto subisce nella bocca è dello stesso ordine di grandezza della larghezza del getto all'altezza del labbro superiore. Di conseguenza il flusso del getto presenta un ampio spettro di armoniche. Se il getto dovesse colpire il labbro in modo esattamente simmetrico, le armoniche pari sarebbero assenti dall'eccitazione. Di solito però il getto è leggermente spostato, e tutte le armoniche sono presenti.

In questa fotografia della bocca di una canna d'organo della famiglia dei principali si vede l'anima munita di piccoli «denti» per ottenere una turbolenza omogenea del getto d'aria. La canna è costruita con una lega di piombo e stagno, particolarmente ricca di stagno, che solidifica dando questa struttura superficiale, se viene fusa in fogli. Al foglio viene poi data la forma della canna.

Come si può facilmente immaginare, la qualità del suono prodotto da una canna aperta è piuttosto diversa da quella di un suono prodotto da una canna tappata. Le frequenze delle armoniche nella forza motrice prodotta dal getto sono multipli interi esatti della frequenza fondamentale dell'onda nel getto. La colonna d'aria nella canna risponderà intensamente a una armonica particolare solo se l'ammettenza acustica della canna è grande, il che corrisponde a un netto picco di risonanza in prossimità della frequenza dell'armonica. Una canna tappata, che presenta solo picchi di risonanza dispari, sopprime tutte le armoniche pari della fondamentale. Il risultato è un caratteristico suono «dolce», in cui le armoniche pari sono deboli, anche se non completamente assenti. Una canna aperta, invece, dà un suono più «brillante» perché risponde a tutte le armoniche della sua frequenza principale.

Le caratteristiche di risonanza di una canna sono in gran parte determinate dalle sue perdite di energia, che possono es-

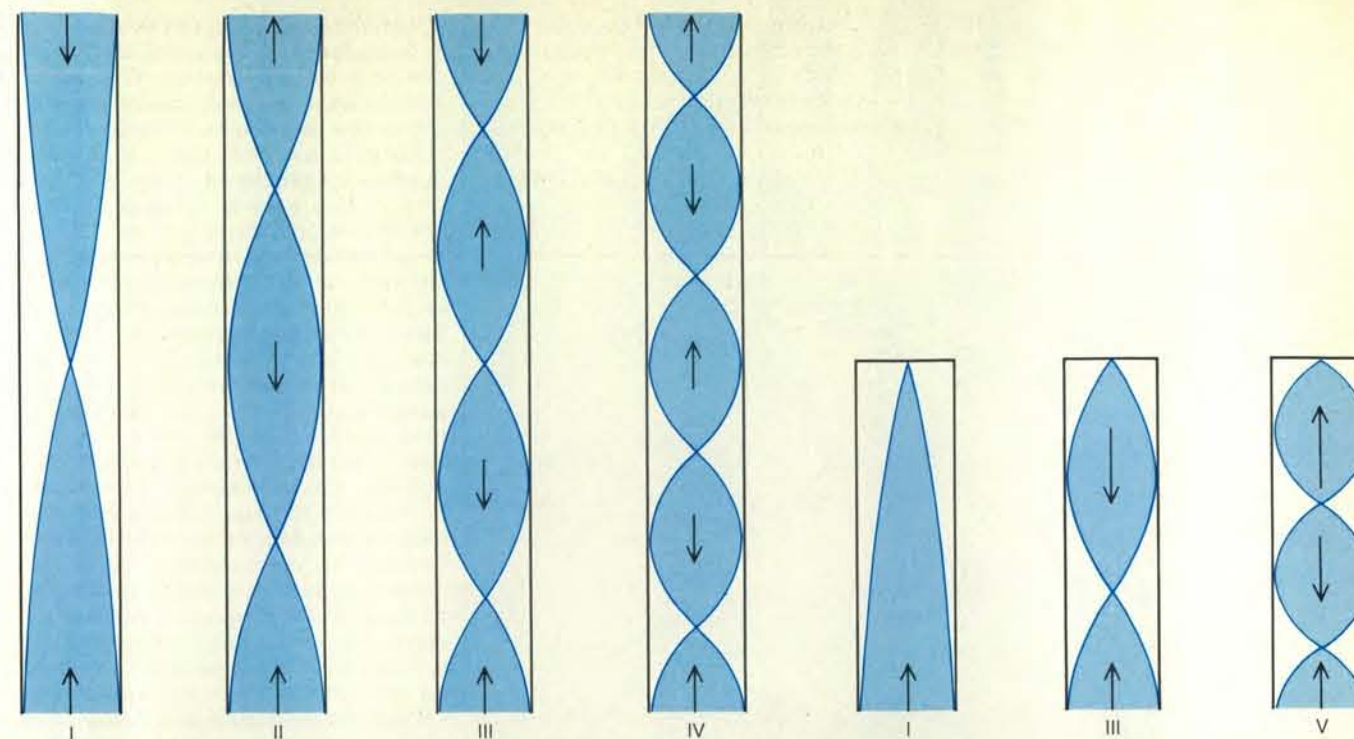
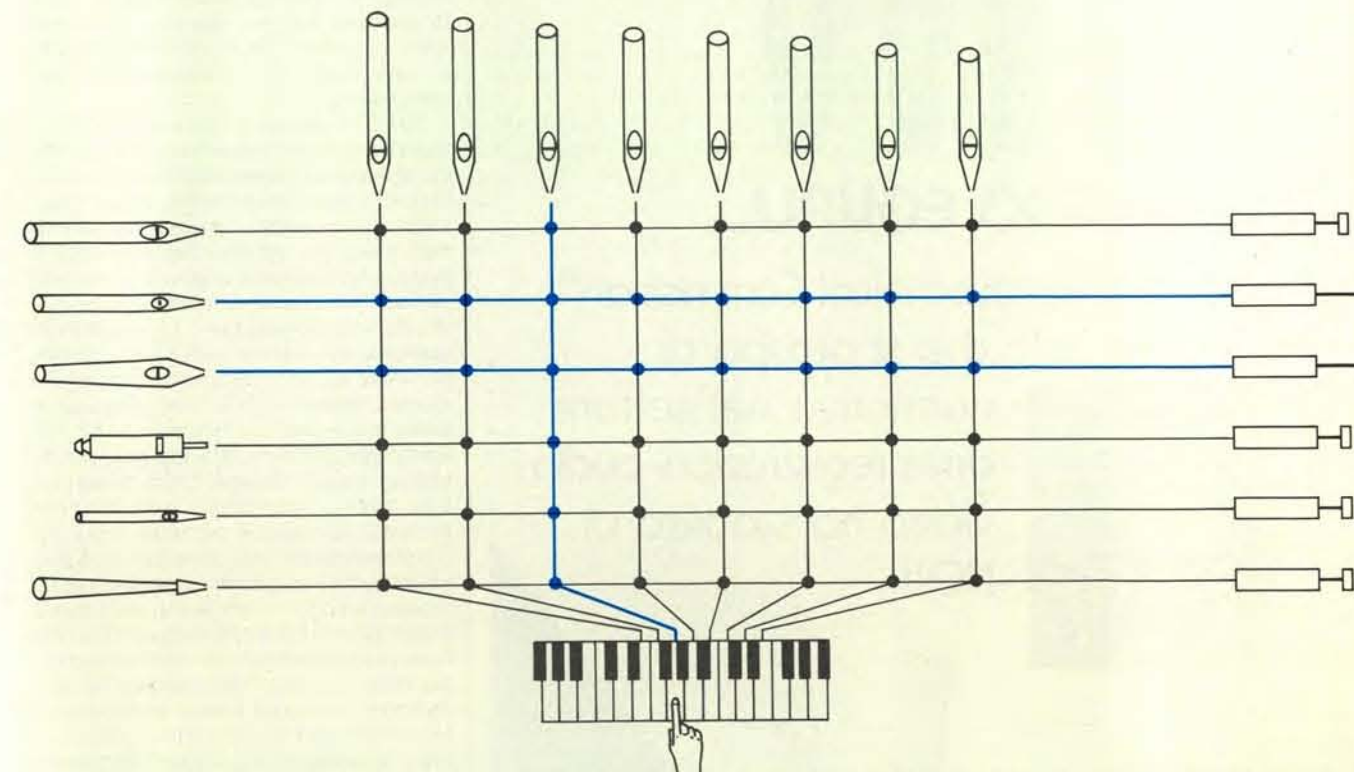
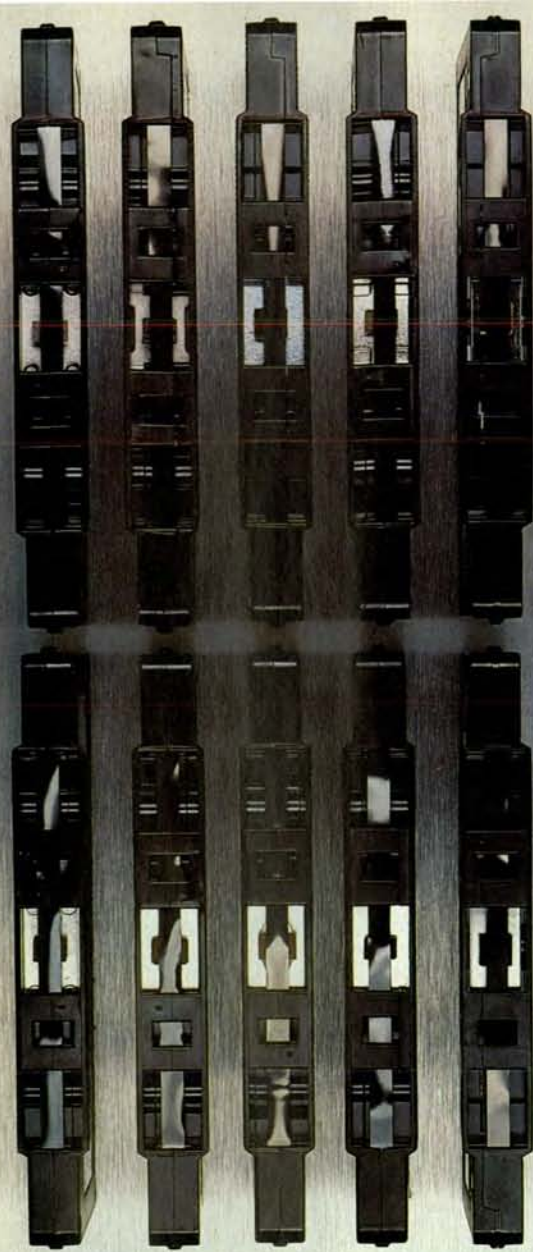
sere di due tipi: 1) perdite per viscosità e per conduzione termica verso le pareti della canna e 2) perdite per irraggiamento dalla bocca e dall'estremità aperta. Le perdite dovute a viscosità e a conduzione termica sono più frequenti per le canne strette che per quelle larghe e si dimostrano più importanti alle basse frequenze che non alle alte. Vale il contrario, invece, per le perdite per irraggiamento.

La conclusione è che, per una data lunghezza di canna e per una data frequenza fondamentale, le canne larghe sono risonatori efficienti e bene intonati solo per la fondamentale e per le prime armoniche, che danno un suono morbido, «flautato». Le canne strette sono buoni risonatori per un'ampia gamma di armoniche e, poiché le frequenze superiori vengono irraggiate con maggiore efficienza delle frequenze più basse, il timbro risulta più simile a quello di uno strumento a corda. Tra questi due estremi stanno le canne dal suono pieno, brillante e argentino che costituisce la caratteristica di un buon organo: sono le canne che costituiscono il registro chiamato comunemente principale o diapason (o anche mostra, prestant, regola primaria, doeff, fondo d'organo, frontispizio, fondamentale).

L'effetto del cambiamento della pressione dell'aria sul suono prodotto, per una canna tipica. Le cifre romane indicano le prime armoniche. Il regime centrale *(in colore)* copre la gamma dei suoni normali, con un buon equilibrio alle normali pressioni dell'aria immessa. A pressioni superiori la canna passa al suo secondo modo proprio di vibrazione; a pressioni basse vi è un secondo modo debole.

L'effetto della posizione del labbro di una canna d'organo sulle componenti armoniche mostra la desiderabilità di uno spostamento del labbro dalla posizione simmetrica. Se il labbro fosse centrato esattamente sul getto, la canna darebbe solamente la frequenza fondamentale *(I)* e la terza armonica *(III)*. Con il labbro spostato nella posizione indicata dalla linea tratteggiata in colore, sono udibili anche la seconda e la quarta armonica, e la qualità timbrica risulta più ricca.

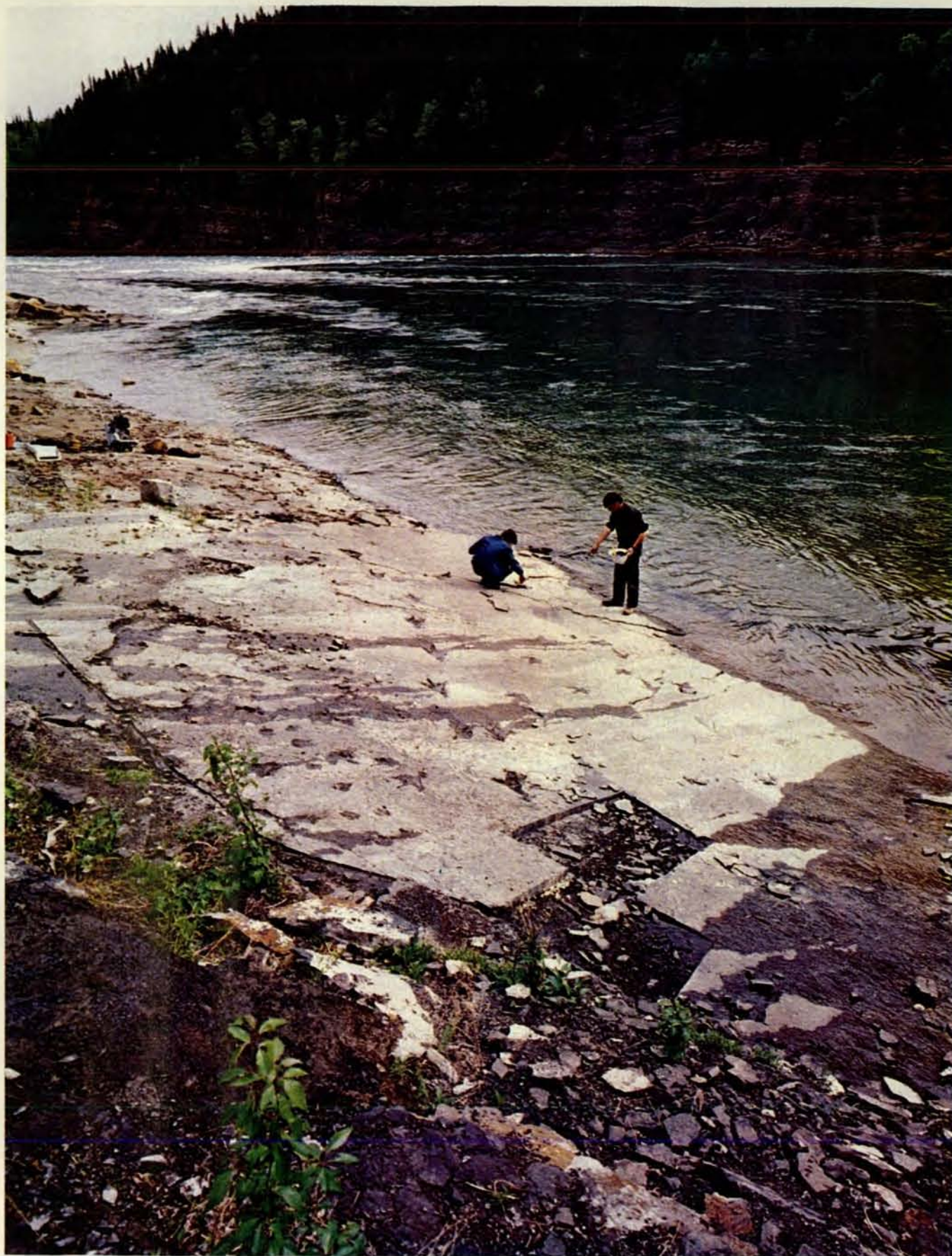
Un organo di grandi dimensioni può avere anche file di canne a corpo conico, con cocuzzolo di copertura munito di un piccolo tubo aperto, o ancora con altre forme geometriche, a cilindro, a prisma, a piramide. Queste caratteristiche sono studiate dai progettisti in modo da modificare le frequenze di risonanza della canna e, in alcuni casi, per intensificare una stretta banda di armoniche superiori in modo da produrre un particolare colore timbrico. Il materiale di cui è fatta la canna riveste un'importanza solo secondaria.

Il fatto che l'aria presente in una canna possa vibrare in uno qualunque di un'ampia serie di modi possibili introduce ulteriori complicazioni nel comportamento acustico delle canne. Per esempio, se la pressione a cui è insufflata l'aria in una canna aperta viene aumentata in modo che nel getto vi sia solo un quarto di lunghezza d'onda del primo modo, f_1, il punto che rappresenta questo primo modo nella spirale dell'ammettenza si sposterà nella metà di destra e il getto cesserà di sostenere questo modo. Al contempo la seconda frequenza di modo proprio, $2f_1$, è tale da corrispondere a mezza lunghezza d'onda nel getto e quindi può essere soggetta a una forza intensa. Il suono della canna pertanto salterà a questo secondo modo, quasi un'ottava al di sopra del primo (la precisa frequenza dipende dalla frequenza di risonanza della canna e dalla pressione dell'aria insufflata).

Un incremento ulteriore della pressione dell'aria può eccitare il modo successivo, $3f_1$, se il lume della canna non è troppo grande. Inversamente, accade spesso che una bassa pressione dell'aria, insufficiente a eccitare il fondamentale, possa eccitare uno dei modi superiori nel secondo giro della spirale dell'ammettenza. Questi suoni particolari, sebbene interessanti in laboratorio, negli organi reali sono sfruttati solamente per effetti speciali.

Il costruttore che ha creato una canna dell'opportuna qualità sonora si trova di fronte al problema di progettare tutta una serie di canne che coprano l'estensione della tastiera e si corrispondano per intensità e contenuto armonico. Un insieme di canne di forma geometrica simile, e semplicemente di dimensioni proporzionalmente maggiori o minori, non soddisfa il requisito: le canne differiscono per quel che riguarda il modo in cui le pareti e le perdite per irraggiamento influenzano le diverse frequenze. Per ottenere un comportamento costante attraverso tutto lo spettro acustico, si deve applicare una regola di scala. Il diametro della canna viene fatto variare in funzione di una opportuna potenza, k (di valore inferiore a 1), della lunghezza della canna. Il risultato è che le lunghe canne dei bassi sono un po' più strette. Il valore di k determinato per via teorica è di 5/6, cioè 0,83, ma se si tiene conto della psicofisica dell'ascolto umano, si dimostra più soddisfacente un valore di 0,75, molto più vicino alla regola empirica sviluppata dai grandi costruttori di organi del XVII e XVIII secolo.

In questo disegno sono rappresentate schematicamente le figure di risonanza a onda stazionaria, tanto per canne aperte quanto per canne chiuse. La larghezza di ogni schema in colore corrisponde all'ampiezza della vibrazione acustica nelle varie parti della canna. Il movimento effettivo dell'aria è parallelo all'asse della canna. Le frecce indicano la direzione del movimento dell'aria nell'arco di una metà del ciclo di vibrazione; durante l'altra metà del ciclo la direzione è invertita. Le cifre romane rappresentano i numeri delle armoniche e sono proporzionali alle frequenze di vibrazione. Una canna aperta ha risonanze per tutte le armoniche della sua fondamentale. Una canna tappata produce la stessa nota con una colonna d'aria alta la metà, ma possiede risonanze solo per le armoniche dispari. Gli schemi di vibrazione si estendono in realtà leggermente al di là delle estremità della canna (motivo per cui sono necessari piccoli aggiustamenti finali) e la complessa geometria della bocca distorce un po' gli schemi all'estremità inferiore della canna, senza tuttavia modificarne la natura «fondamentale».

Le canne di un organo sono disposte come le righe e le colonne di una matrice. In questo diagramma semplificato ciascuna riga (quella che viene chiamata una fila di canne) è costituita da canne tutte dello stesso tipo: a ogni canna corrisponde una nota *(in alto)*. Ogni colonna, che è associata a una nota della tastiera *(in basso)*, fornisce l'accesso a tutte le canne dei diversi tipi *(a sinistra)*. Le placchette sulla consolle dell'organo *(a destra)* consentono, se inserite, il passaggio dell'aria verso le canne di una data fila; i tasti della tastiera, se schiacciati, consentono il passaggio dell'aria a tutte le canne che si trovano sul canale corrispondente alla loro nota. Il meccanismo è organizzato in modo che l'aria raggiunga una canna solo se sono attivate sia la riga, sia la colonna relativa. Nella situazione qui raffigurata, suoneranno solo due canne.

Per finire viene il meccanismo attraverso il quale viene controllato il suono della grande quantità di canne in un organo di grandi dimensioni (meccanismo che è di particolare importanza, ovviamente, dal punto di vista dell'esecutore). Il progetto di fondo è deliziosamente semplice e concettualmente è simile all'organizzazione per righe e colonne di una matrice. Le canne sono disposte in file, che corrispondono alle righe della matrice: tutte le canne che si trovano in una medesima fila hanno la stessa qualità timbrica, e la fila è costituita da una canna per ciascuna nota della tastiera o della pedaliera. Il passaggio dell'aria verso ciascuna fila è controllato da una placchetta su cui è riportato il nome della fila (il registro). Il passaggio dell'aria verso le canne associate a ciascuna nota, cioè verso una colonna della matrice, è controllato dal tasto della tastiera (o dal pedale della pedaliera) corrispondente a quella nota. Una data canna suonerà, quindi, se e solo se si trova in un registro che è stato inserito mediante l'apposita placchetta e su un canale di tasto il cui tasto corrispondente è stato premuto.

Oggi possiamo immaginare moltissimi modi di realizzare questo schema, mediante logica digitale e valvole azionate elettricamente al di sotto di ciascuna canna. I primi organi possedevano semplici valvole e leve meccaniche, per far passare l'aria nei canali dei tasti, e liste perforate mobili per controllare l'ingresso dell'aria in una fila completa di canne. Oltre ad avere il pregio della semplicità e della durata, il sistema meccanico dava all'esecutore un controllo preciso sulla velocità di apertura delle valvole dei canali e un senso di partecipazione al funzionamento di uno strumento altrimenti piuttosto meccanico.

Nel XIX secolo e agli inizi del XX, i grandi organi sono stati dotati di ogni tipo di meccanismo elettromeccanico ed elettropneumatico, ma negli ultimi tempi l'accento è caduto ancora sull'azione meccanica per quanto riguarda tasti e pedali, con raffinati controlli elettronici per la formazione di combinazioni di registri durante l'esecuzione. L'organo della sala concerti della Sydney Opera House, per esempio, è il più grande organo del mondo dotato di azione meccanica. Completato nel 1979, possiede 10 500 canne distribuite in 205 file, con cinque tastiere e una pedaliera. L'azione dei tasti è meccanica, ma è duplicata da un'azione elettrica, cosicché si possono usare accoppiatori elettrici. In questo modo è possibile registrare in forma digitale una esecuzione «dal vivo» per poi usare il nastro registrato per azionare l'organo e ricreare l'esecuzione originale. I registri e i pistoncini delle combinazioni sono tutti elettrici o elettropneumatici e sono controllati da microelaboratori a programmi memorizzati, di notevole versatilità. In questo modo le migliori risorse della tecnologia moderna, unite alla progettazione tradizionale e alla abilità dei singoli artigiani che seguono principi noti da secoli, creano il suono flessibile e maestoso di un grande organo.

Un ricco giacimento di orme fossili, portato alla luce nel corso di cinque anni di lavoro febbrile nel Peace River Canyon, nella British Columbia, prima che il sito venisse allagato per la costruzione di una diga, comprendeva questa sequenza continua di impronte di dinosauro conservate in sedimenti del Cretaceo inferiore. In primo piano, uno dei blocchi rettangolari con sopra impressa un'orma è stato rimosso. In secondo piano, alcuni ricercatori del Natural History Museum della provincia di Alberta stanno preparando il calco di altre orme. In meno di sette chilometri il personale addetto al lavoro di ricupero ha rinvenuto 1700 orme, che rappresentano almeno dieci differenti specie di dinosauri.

Orme di animali estinti

Sin dalla loro prima comparsa sulla terraferma, 370 milioni di anni fa, i vertebrati hanno lasciato impronte in molti sedimenti. La maggior parte delle specie estinte a noi note lo sono soltanto in virtù delle loro orme

di David J. Mossman e William A. S. Sarjeant

Gli scheletri di animali terrestri estinti che si possono vedere in mostra nei musei fanno pensare che le principali testimonianze fossili relative a questi animali siano le ossa, ma non è così. Le ossa degli animali sono molto più rare delle loro tracce. Anzi, molti animali terrestri estinti ci sono noti unicamente per le loro tracce; pare che i resti dei loro corpi non si siano mai conservati. In Inghilterra, per esempio, i sedimenti del Triassico medio del Worcestershire e quelli del Triassico superiore del Nottinghamshire, gli uni e gli altri depositatisi nell'«Era dei rettili» mesozoica, non contengono fossili di vertebrati. Le orme rinvenute in queste stesse formazioni indicano però che almeno otto diverse specie di rettili calpestarono i sedimenti del Worcestershire e altre sei fecero la stessa cosa nel Nottinghamshire.

Anche la più antica invasione nota della terraferma da parte di vertebrati è analogamente testimoniata dalle tracce che si trovano nella formazione Old Red Sandstone delle isole Orkney, risalente al Devoniano medio (da 370 a 360 milioni di anni fa). Queste tracce testimoniano il passaggio di un pesce primitivo che, al pari degli odierni dipnoi, viveva alla mercé dei periodi stagionali di tempo asciutto. Questo pesce lasciò la propria traccia strisciando col ventre sulla terraferma per andare da uno specchio d'acqua all'altro, servendosi per avanzare delle sue tozze pinne lobate.

Anche se il Devoniano può sembrare un periodo molto remoto a noi che viviamo oggi, non lo è poi così tanto se viene considerato nella scala temporale assoluta della storia della Terra. La valutazione è facile se i 4,5 miliardi di anni trascorsi da quando la Terra si formò vengono considerati come un solo anno, in cui ogni giorno abbia la durata di 12,3 milioni di anni. A una scala temporale di questo genere, le prime forme di vita terrestre - piante primitive simili alle attuali alghe unicellulari - comparvero nei mari ai primi di maggio. Le forme di vita pluricellulari però non apparvero che ai primi di novembre.

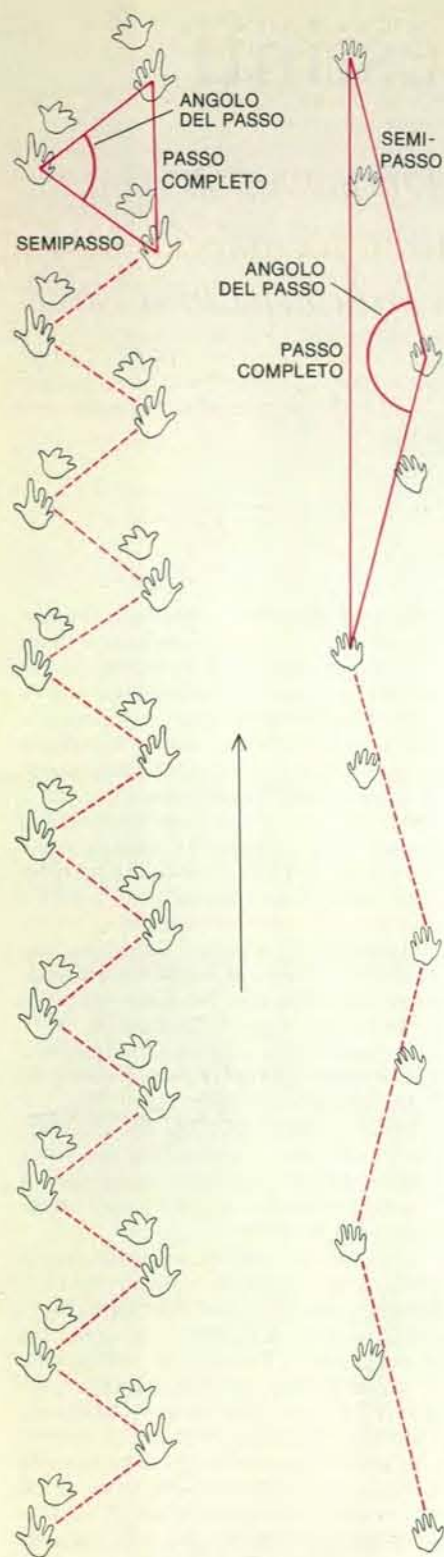
Verso il 20 novembre pesci primitivi nuotavano nelle acque del pianeta. Verso la fine del mese certi loro discendenti si avventurarono sulla terraferma. Il 7 dicembre i rettili erano diventati ormai gli animali predominanti sulla terraferma e alla metà di dicembre comparvero i primi mammiferi. Verso le cinque del pomeriggio dell'ultimo giorno dell'anno due ominidi primitivi lasciarono le proprie orme su uno strato fresco di cenere vulcanica nella piana di Laetoli, nel Kenya (si veda l'articolo *Le orme fossili di Laetoli* di R. L. Hay e M. D. Leakey, in «Le Scienze» n. 164, aprile 1982). Il genere al quale noi apparteniamo, *Homo*, comparve soltanto un'ora, minuto più minuto meno, prima di mezzanotte (cioè circa 500 000 anni fa). Così l'intero arco della presenza dei vertebrati sulla terraferma occupa meno di sei settimane dell'«anno» di storia della Terra. Anche se abbraccia circa 360 milioni di anni, questo importante intervallo di tempo è pur sempre una parte molto piccola della storia del pianeta.

Documenti terrestri

In tutto quest'arco di tempo si sono conservati due tipi di testimonianze che aiutano i paleontologi a interpretare la vita animale del passato. Uno è costituito dai resti degli animali stessi: ossa fossilizzate, articolate o separate, con o senza strutture associate come denti, squame o penne e, in rari casi, con tracce fossilizzate di pelle o di altri tessuti. Da questi resti è possibile ricostruire con buona approssimazione l'aspetto dell'animale e dedurre il modo in cui esso si spostava e si nutriva. Un animale però ha soltanto un corpo, e i suoi resti, ammesso che si conservino, si possono trovare in un solo luogo. L'altro tipo di testimonianza è rappresentato dalle tracce e dalle orme dell'animale, le quali costituiscono una indicazione molto più estesa e dinamica. Studiando questi indizi del comportamento di un animale, i paleontologi sono in grado, per così dire, di vedere l'animale in azione.

I crossotterigi che per primi si avventurarono con le loro pinne lobate sulla terraferma in tempi devoniani diedero origine agli anfibi a quattro zampe. In termini evolutivi la transizione fu rapida: le più antiche impronte di anfibi si trovano in formazioni del Devoniano superiore. La traccia formata da una sequenza continua di impronte di anfibi è caratteristicamente larga e mostra quel passo corto che è tipico di una locomozione quadrupede piuttosto inefficiente. Le singole orme presentano il segno pieno della pianta del piede (che denota un'andatura lenta) e dita prive di artigli che si aprono a ventaglio e sono dirette verso l'esterno anziché in avanti. Le orme di anfibi sono abbondanti nei sedimenti del Carbonifero, il periodo che seguì il Devoniano. Sono meno abbondanti nel Permiano (da 280 a 225 milioni di anni fa) e sono rare nella documentazione fossile successiva, soprattutto a causa del fatto che gli anfibi che sopravvissero preferirono, così come fanno tuttora, condizioni ambientali di presenza d'acqua, non adatte alla conservazione di impronte.

L'ascesa dei rettili a una posizione di predominio si nota chiaramente nel Carbonifero superiore e nel Permiano via via che il numero delle loro orme aumenta costantemente. Tanto la diversificazione di questi animali quanto la sempre maggior efficienza della loro locomozione sono rese evidenti da mutamenti intervenuti nella forma delle dita, nonché dallo sviluppo di un quinto dito simile a un pollice e dalla comparsa di artigli. In alcuni gruppi di rettili la traccia formata da una sequenza continua di impronte si restringe e il passo si allunga. Gli animali si muovevano più velocemente ma la loro andatura rimaneva quadrupede. Altre impronte indicano che nel Permiano superiore alcuni rettili di piccole dimensioni mantenevano l'andatura quadrupede quando si muovevano adagio ma ne adottavano una bipede quando si muovevano in fretta. Il loro corpo faceva perno sulle anche, e una lunga coda faceva da contrappeso agli arti anteriori e al tronco quasi eretto.

Questa fu la linea di rettili che diede origine ai dinosauri all'inizio dell'Era

Raffigurazione schematica di sequenze continue di impronte. A sinistra, la sequenza larga caratteristica di un anfibio; vi sono raffigurate le orme consecutive delle zampe anteriori e posteriori di uno degli anfibi estintisi verso la fine dell'Era paleozoica. Come indica il triangolo quasi equilatero in colore, tanto il passo completo quanto il semipasso dell'animale erano brevi. La sequenza più stretta rappresentata a destra, caratteristica di un rettile, presenta le orme consecutive degli arti anteriori e posteriori. Tanto la distanza coperta dal passo completo quanto quella coperta dal semipasso sono proporzionatamente maggiori di quelle dell'anfibio, segno di una locomozione più efficiente.

mesozoica, 225 milioni di anni fa. Molti gruppi di dinosauri triassici erano in gran parte, o interamente, bipedi. Anche nei gruppi che più tardi ritornarono alla locomozione quadrupede vi era una netta differenza nelle dimensioni degli arti. Il corpo continuava a far perno sul bacino, e i piedi posteriori di solito erano più grandi di quelli anteriori. Di tutte le orme fossili, quelle dei dinosauri, che si trovano in gran copia nei sedimenti terrestri dell'Era mesozoica in quasi tutte le parti del mondo, sono le più spettacolari.

Col tempo, in due gruppi di dinosauri di piccole dimensioni si svilupparono delle ali. In uno di questi gruppi, quello dei pterosauri, si trattava di membrane che permettevano agli animali un volo planato aprendosi verso l'esterno dai fianchi del corpo, sostenute dal quarto dito, di forma allungata, dell'arto anteriore. W. L. Stokes, dell'Università dello Utah, ha scoperto tracce fossili di questi rettili volanti in sedimenti della formazione Morrison, un deposito del Giurassico superiore dell'Arizona. La sequenza delle impronte rende chiaro che quegli animali non solo erano capaci di planare, ma riuscivano anche a camminare agevolmente a quattro zampe, con la membrana alare che veniva trascinata per terra dietro gli arti anteriori.

Nel secondo gruppo una modificazione delle squame epidermiche portò alla formazione di penne e a un modo di volare più efficiente. L'evoluzione degli uccelli dai dinosauri non produsse alcuna modificazione immediata delle zampe posteriori, ma non tardarono ad apparire orme di tipo nettamente aviario. La transizione è evidente nelle orme di uccelli rinvenute in magnifico stato di conservazione nei sedimenti del medio Cretaceo della British Columbia orientale da Philip J. Currie, del Tyrrel Museum of Paleontology di Drumheller, nella provincia canadese di Alberta.

Non tutti i rettili mesozoici diventarono bipedi. I sinapsidi, rettili che per tutta la loro storia rimasero sempre separati dai dinosauri, furono uno dei gruppi che conservarono costantemente una locomozione quadrupede. Anche qui però il passo degli animali si allunga, la sequenza continua di impronte si restringe e le dita sono dirette in avanti: tutti cambiamenti che denotano una locomozione sempre più veloce ed efficiente. La struttura del piede dei sinapsidi diventò sempre più simile a quella del piede di un mammifero. A dire il vero, nei sedimenti del Triassico superiore e del Giurassico inferiore della provincia di São Paulo il paleontologo brasiliano Giuseppe Leonardi ha rinvenuto numerose orme di forma così esattamente intermedia tra quelle dei rettili e quelle dei mammiferi da poter essere attribuite indifferentemente all'uno o all'altro gruppo.

Con la fine del Mesozoico le impronte dei dinosauri scompaiono dalla storia geologica. Dall'inizio dell'Era cenozoica, 65 milioni di anni fa, la varietà delle impronte di mammiferi corrisponde alla sempre maggiore diversificazione di questi vertebrati evoluti, diversificazione che è uno degli avvenimenti evolutivi più importanti del Terziario, la prima e più lunga parte del Cenozoico. Sembra logico pertanto che con il Quaternario, via via che la storia geologica si avvicina al presente, si trovi, fra le impronte di mammiferi di Laetoli, la prova del fatto che nella documentazione fossile siano entrati alcuni dei precursori a stazione eretta dell'uomo.

La prima scoperta

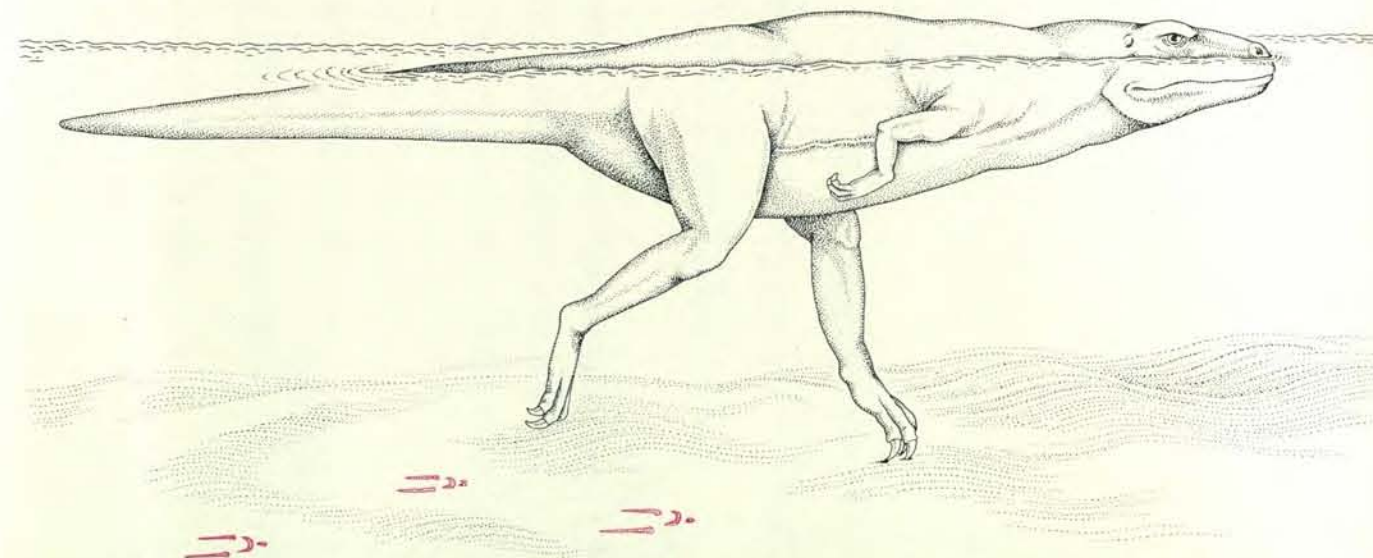
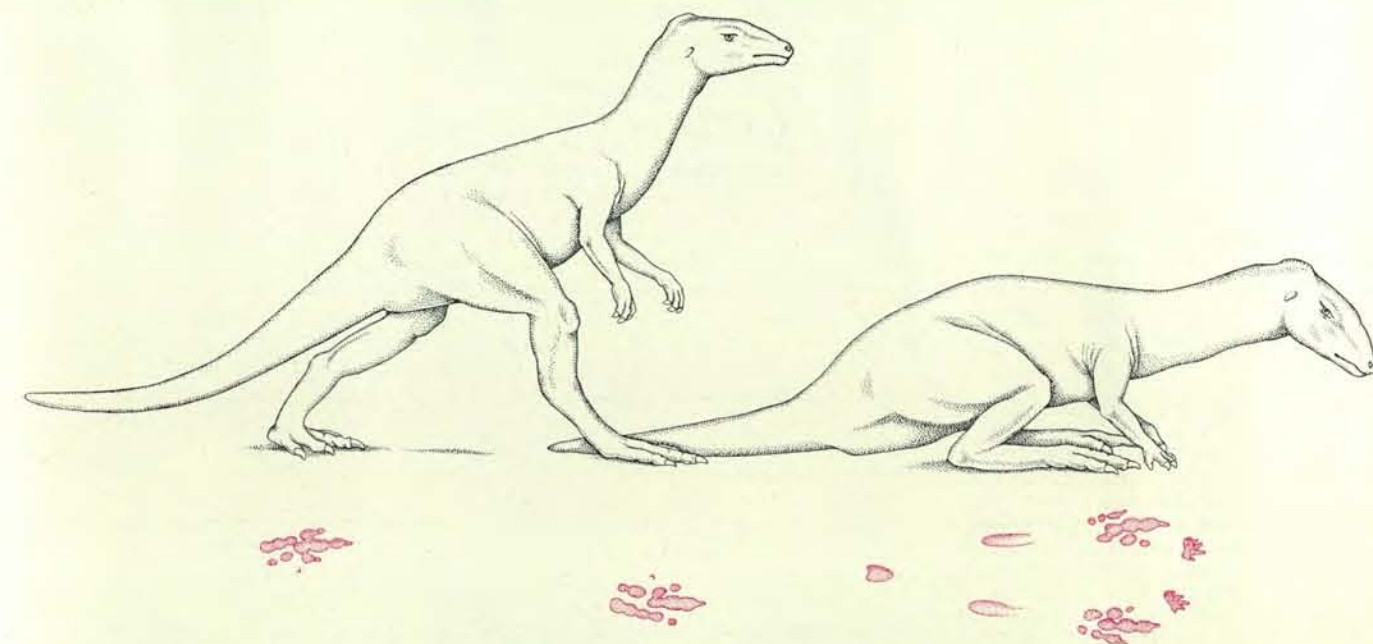
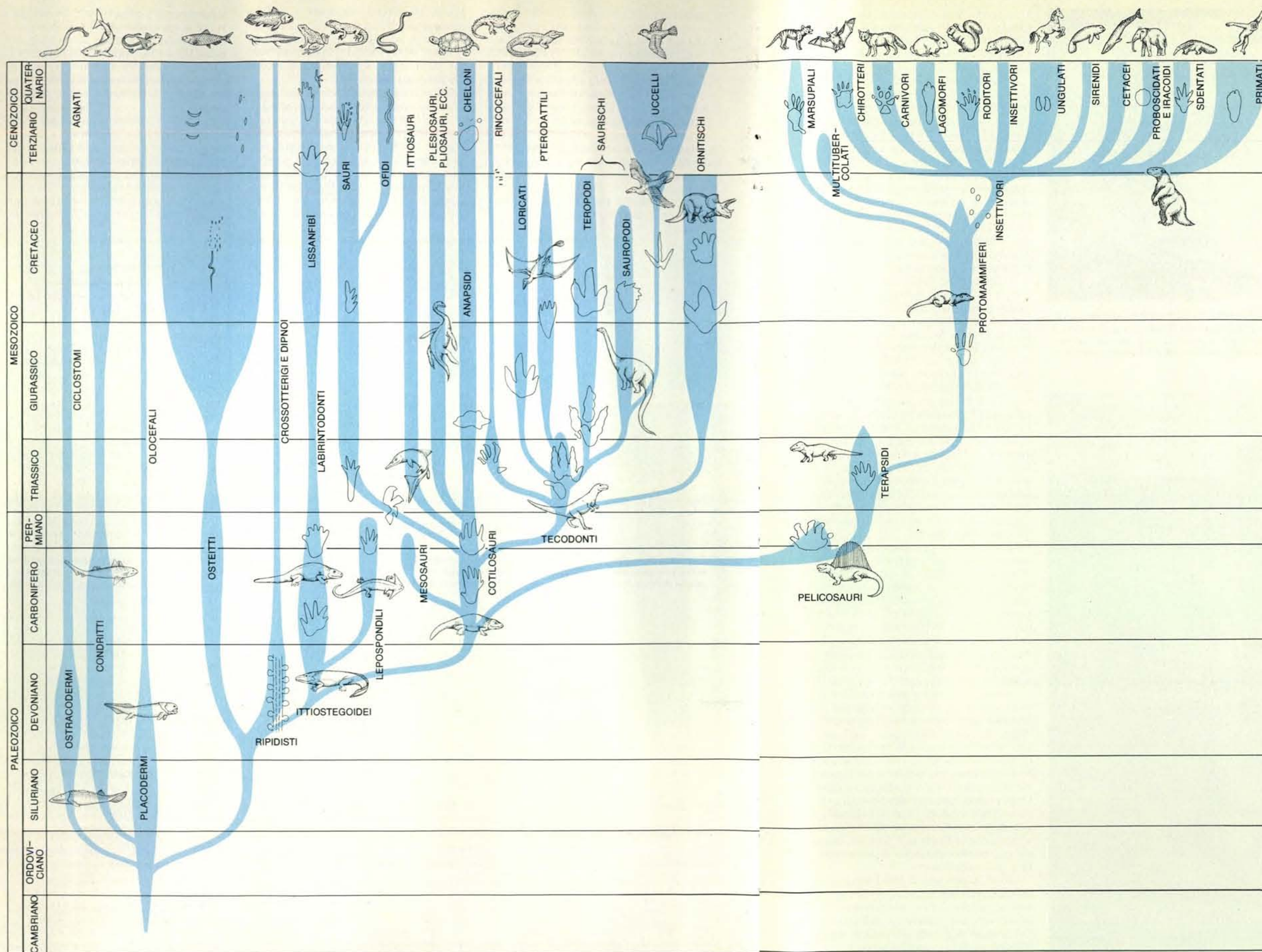
Qual è la più antica scoperta documentata di orme fossili di vertebrati? Si potrebbe pensare che sia stata fatta in Europa da uno dei pionieri della geologia, ma non è così. La località fu la valle del fiume Connecticut nei pressi di South Hadley, nel Massachusetts, dove sono affiorati strati di arenaria rossa che datano dal Triassico superiore al Giurassico inferiore. Lo scopritore fu Pliny Moody, un bracciante agricolo del luogo, l'anno fu il 1802 e le impronte erano quelle di dinosauri di piccole dimensioni. A quel tempo però i dinosauri erano del tutto sconosciuti, e le orme furono scambiate per quelle di grossi uccelli. Taluni arrivarono addirittura al punto di avanzare l'ipotesi che si trattasse delle impronte del corvo di Noè.

Passò un quarto di secolo prima che fosse pubblicata una relazione scientifica riguardante orme fossili. Si trattava di orme trovate in una formazione sedimentaria della Scozia, la New Red Sandstone. A quel tempo l'età della formazione non era nota con precisione, ma oggi si sa che quei sedimenti si accumularono nella parte finale dell'Era paleozoica. Lo scopritore fu un ecclesiastico e naturalista dilettante del luogo, il reverendo Henry Duncan. La sua descrizione delle impronte, pubblicata nel 1828, riempì di stupore il mondo scientifico dell'epoca.

Le orme, che mostrano il movimento di un animale con un passo corto e uno scarto notevole fra le zampe di destra e quelle di sinistra, sono ora riconosciute come quelle di un caseasauro, un animale appartenente a un gruppo di rettili erbivori grandi e massicci. Al pari dei dinosauri della valle del fiume Connecticut, i caseasauri erano ancora sconosciuti, e così le impronte furono attribuite, del tutto ragionevolmente, a una qualche specie di tartaruga acquatica. Venuto a conoscenza della scoperta di Duncan, un eminente geologo inglese, William Buckland, decise di verificare l'ipotesi dell'impronta di tartaruga. John Murray III, un famoso editore, così descrisse l'esperimento.

«Sabato scorso mi sono recato a una riunione in casa di Mr. Murchison, organizzata per osservare delle tartarughe acquatiche nell'atto di camminare su una pasta. Il professor Buckland fungeva da maestro di cerimonia. Erano presenti molti altri geologi e uomini dotti, fra i quali il dottor Wollaston. Da principio gli animali decisero di fare i difficili e di rimanere immobili. A un certo punto, per farli muovere, si fece appello all'ingegnosità del professore, il quale ci provò dando parecchi colpetti con le mani sul posteriore delle tartarughe, ma non c'era verso che queste si muovessero. E ne avevano ben donde: provando a sollevarle, infatti, si scoprì che le tartarughe si erano attaccate così rapidamente alla pasta da poter essere rimosse soltanto portandosene dietro una mezza libbra attaccata a ogni zampa. Si dovette dar mano a un matterello per lavorare di nuovo la pasta, e le dita geologiche non disdegnarono queste incombenze culinarie. Era proprio una scena magnifica vedere tutti quei sapienti, mezzo infarinati, lavorare alacremente con le maniche rimboccate. Alla fine, sono lieto di poterlo dire, i loro sforzi furono coronati da successo: si riuscì a ottenere una pasta di consistenza appropriata, e gli animali ci camminarono sopra in modo molto soddisfacente, tant'è vero che molti, venuti lì per schernire, furono più propensi a credere.»

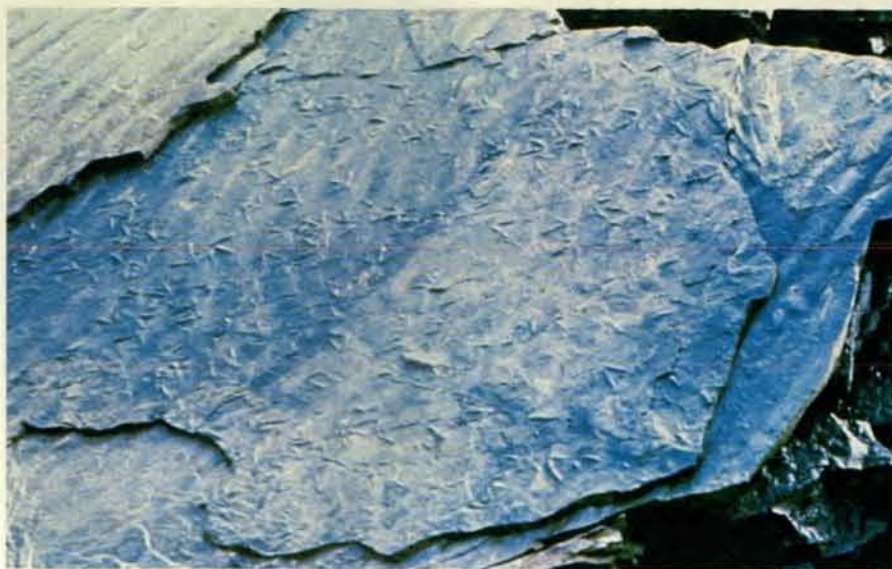
Dopo la scoperta di Duncan molte altre impronte cominciarono a essere segnalate, ma non necessariamente accettate. Nel 1841, per esempio, sulla costa sotto Horton Bluff, nella Nuova Scozia, William Logan, un turista scozzese, trovò in uno strato di roccia sedimentaria orme ben conservate di vertebrati. La formazione risaliva al Carbonifero, che però, secondo l'opinione prevalente ai tempi di Logan, era un periodo troppo recente della storia geologica perché vi potessero essere esistiti vertebrati al di fuori dei pesci. Quando, tornato in Inghilterra, Logan presentò il proprio reperto ai membri della Geological Society, essi si rifiutarono di credere a quella testimonianza.

Logan fu vendicato soltanto nel 1872. In quell'anno John W. Dawson, che lavorava con uno dei padri fondatori della geologia, Sir Charles Lyell, stava studiando alcuni strati del Carbonifero sulla costa orientale della baia di Fundy. Gli strati in questione comprendevano i giacimenti di carbone di Joggins nella Nuova Scozia,

Impronte di un dinosauro carnivoro rinvenute recentemente in sedimenti giurassici mostrano chiari segni di dita e di artigli che non danno alcuna indicazione di una pressione del piede sul fondo. La raffigurazione del carnivoro nell'atto di nuotare si basa sulla ricostruzione di un *Megalosaurus* fatta da Matthew Hyman. Le impronte indicano che i dinosauri carnivori non vivevano affatto confinati sulla terraferma.

Impronte di dinosauro scoperte nelle arenarie rosse della valle del fiume Connecticut indicano che l'animale che le ha lasciate era fondamentalmente bipede, come si vede a sinistra, ma che a volte camminava a quattro zampe, come è mostrato a destra. Si tratta di uno dei tanti rettili estinti noti soltanto dalle loro impronte; da queste Richard Lull, della Yale University, ha potuto ricostruire l'ornitopodo qui raffigurato.

L'evoluzione dei vertebrati ha fornito alla documentazione fossile numerose impronte, fra cui la sequenza terrestre più antica è quella di un pesce crossotterigio lasciata nella Old Red Sandstone delle isole Orkney ai tempi del Devoniano medio *(quinta banda da sinistra)*. In seguito, come risulta dalle orme tipiche, hanno lasciato le proprie impronte tutta una serie di anfibi, rettili e mammiferi. L'ultima impronta, a destra in alto, è quella di un ominide nel Kenya.

dove si erano conservati come fossili i ceppi di numerosi alberi. Dawson aveva previsto che nell'interno cavo di questi ceppi si sarebbero potuti trovare resti fossili di anfibi e rettili di piccole dimensioni. La sua previsione si dimostrò esatta: all'interno dei ceppi furono rinvenuti effettivamente resti di rettili. Forse gli animali erano rimasti intrappolati nelle cavità mentre cercavano acqua o insetti. In ogni caso, era chiaro ormai che la fauna del Carbonifero comprendeva non solo pesci, ma anche vertebrati terrestri.

Interesse in declino

Mentre nella seconda metà dell'Ottocento vennero pubblicate molte descrizioni di impronte fossili di vertebrati, nella prima metà del Novecento gli scritti sull'argomento videro la luce in numero sempre minore. Questo ramo della paleontologia, detto icnologia (dal greco *ikhnos*, traccia), cadde progressivamente in disgrazia presso i geologi. Le ragioni non sono difficili da trovare. Tra il 1930 e il 1960 i pochi scritti in materia erano opera soprattutto di dilettanti anziché di professionisti, uscivano su pubblicazioni relativamente sconosciute ed erano di livello qualitativo scadente. Le descrizioni che gli autori facevano delle impronte che avevano scoperto erano deplorevolmente insufficienti, e ben di rado i loro reperti venivano messi in relazione in maniera sistematica con impronte che erano state descritte in precedenza. Le fotografie delle impronte, ammesso e non concesso che venissero accluse, erano fatte in genere con illuminazione dall'alto anziché con l'illuminazione laterale necessaria per mettere in evidenza le impronte poco profonde o il rilievo limitato dei calchi. Il più delle volte le illustrazioni non erano neppure costituite da fotografie scadenti, ma da disegni al tratto del tutto inadeguati. Quasi mai gli autori indicavano dove i reperti avrebbero potuto essere esaminati da altri, soprattutto perché la maggior parte dei reperti stessi non erano mai sistemati in collezioni di studio. Qui la colpa non era necessariamente degli autori. Essendo grandi e scomode da maneggiare, le lastre con le impronte tendono a non essere molto popolari presso i conservatori dei musei. Una volta pervenute a destinazione, le lastre finiscono con tutta probabilità fuori dalle mura del museo, stese a terra, dove vengono consumate dai piedi dei visitatori, oppure addossate a un muro, dove l'esposizione agli agenti atmosferici alla fine le distrugge.

Negli anni cinquanta le ricerche svolte dal paleontologo tedesco Walter H. Häntzschel e gli sforzi missionari del suo allievo Adolf Seilacher ebbero come risultato un risveglio dell'interesse per le impronte fossili lasciate da organismi marini. Anche allora però non si ebbe un equivalente risveglio dell'interesse per le orme di vertebrati. Di queste oggi si occupa solo un piccolo numero di ricercatori, fra i quali in particolare Giuseppe Leonardi e Rodolfo Casaniquela in America del Sud, Donald Baird alla Princeton

Orme di uccelli compaiono qua e là in sedimenti mesozoici della British Columbia, nel famoso sito del Peace River Canyon, dove nel 1922 fu rinvenuto un gran numero di impronte di dinosauri. I caradriidi che lasciarono queste impronte sono altrimenti sconosciuti nella documentazione fossile. Le loro orme sono le uniche tracce di uccelli rinvenute in strati del Cretaceo inferiore.

Questa sequenza di impronte lasciate da un mammifero xenartro, grandi quanto quelle di un elefante, attraversa un affioramento di una formazione geologica dell'Argentina: quella del Rio Negro. Si tratta di impronte che risalgono al Pliocene superiore, tre milioni e mezzo di anni fa.

University, Justin Delair in Inghilterra, Georges Demathieu e Albert F. de Lapparent in Francia, Hartmut Haubold in Germania Orientale e O. S. Vialov in Unione Sovietica.

Oltre ai problemi di un'inadeguata letteratura scientifica, lo studioso di orme fossili si trova di fronte ad altre difficoltà. Perché si formino delle orme nette è necessario innanzi tutto che ci sia uno strato di sedimento coesivo, umido e a grana fine. Poi bisogna che l'animale proceda su di esso lentamente, lasciando l'impronta particolareggiata dei piedi anteriori e posteriori (ammesso che si tratti di un quadrupede). In queste condizioni ideali potrebbero conservarsi perfino i contorni esatti degli artigli o delle unghie, la forma dei cuscinetti carnosi e la struttura delle squame.

Di orme belle come queste in realtà non se ne trovano molte. Se è a grana troppo grossa, il sedimento non conserva i particolari. Se è troppo bagnato, le cavità più profonde si riempiono d'acqua, alterando la forma dell'orma. Se è troppo asciutto, non è abbastanza coesivo da conservare l'impronta. Non è escluso poi che l'orma venga cancellata da un forte vento, e se è stata lasciata su una spiaggia l'alta marea potrebbe fare la stessa cosa. Perfino la deposizione di sedimento sopra le impronte, che è chiaramente fondamentale per la loro conservazione, potrebbe significarne la cancellazione. Se il nuovo sedimento è troppo simile a quello sul quale è stata lasciata l'impronta, i due strati tenderanno a non lasciarsi separare, e così non è da escludere che le orme non vengano mai scoperte.

Le condizioni più favorevoli per la conservazione delle impronte si verificano quando acque alte si sono ritirate facendo affiorare un nuovo sedimento a grana fine. Sulla riva del mare, questo significa dopo il livello più alto di una marea sizigiale. Nell'entroterra, significa dopo una stagione piovosa che ha innalzato al suo livello massimo un corso o una raccolta d'acqua. In una superficie umida di questo genere le impronte si solidificano a mano a mano che il sedimento si asciuga e viene sepolto dal sedimento che si deposita successivamente. Anche in questo caso, però, può darsi che non sia l'impronta a conservarsi quanto piuttosto il suo calco, soprattutto se il sedimento più recente è più coesivo e a grana più grossa di quello precedente. Questi calchi devono essere cercati nella parte inferiore degli strati sedimentari e potrebbero anche sfuggire all'attenzione, a meno che lo strato non sia stato capovolto, poniamo, dal crollo di una scogliera o da lavori di scavo.

Infine, la probabilità che l'orma di un animale terrestre si conservi e venga in seguito riconosciuta è direttamente proporzionale alla grandezza dell'animale. Gli animali di grandi dimensioni hanno bisogno sia di muoversi in lungo e in largo per cercare il cibo sia di avvicinarsi spesso alle pozze d'acqua. Inoltre, le loro orme si imprimono più profondamente e quindi si distruggono meno rapidamente.

Pur essendo molto più numerosi di quelli grandi, gli animali di piccole dimensioni non coprono nel loro girovagare una vasta area, hanno bisogno di una minore quantità di cibo e di acqua e lasciano orme meno profonde. Ecco perché è più probabile che si conservino (e molto più probabile che vengano riconosciute) le impronte di un dinosauro che non quelle di una lucertola.

Analisi delle impronte

Da un punto di vista ideale, una sequenza continua di impronte adatta per un'analisi dovrebbe consistere di almeno tre orme o calchi. Nei quadrupedi la locomozione ha inizio con il piede posteriore di un lato, seguito dal piede anteriore dello stesso lato e poi dal piede posteriore e da quello anteriore dell'altro lato. Nella locomozione veloce due o tre piedi toccano terra simultaneamente; in quella più lenta toccano terra simultaneamente tre piedi o tutti e quattro. Quando l'animale procede a balzi tutte e quattro le orme si trovano una vicina all'altra senza mai sovrapporsi; impronte di questo genere sono estremamente rare nella documentazione fossile, anche se alcune sono state rinvenute di recente da Leonardi e da uno di noi (Sarjeant) in una formazione brasiliana del Mesozoico. Di norma le tracce lasciate da bipedi presentano alternativamente il piede destro e quello sinistro; ben di rado le orme sono fianco a fianco.

Il primo compito di chi si occupa di analisi delle orme è quello di prendere quattro misure fondamentali di una loro sequenza. La prima è la misura del passo completo, presa da un punto stabilito di un'orma allo stesso punto dell'orma successiva dello stesso piede. La seconda è la misura del semipasso, cioè della distanza tra l'orma del piede anteriore destro e quella del piede anteriore sinistro e tra l'orma del piede posteriore destro e quella del piede posteriore sinistro; di solito la lunghezza del passo completo è identica per le zampe anteriori e per quelle posteriori, ma nel semipasso si possono anche riscontrare differenze notevoli. La terza è la misura dell'angolo del passo, cioè dell'angolo che si individua unendo il punto medio di tre orme successive delle zampe posteriori o di tre orme successive delle zampe anteriori. La quarta è la misura della larghezza della sequenza delle impronte. Come si è già notato, la caratteristica di un animale che cammina in maniera poco efficiente è una sequenza larga e un passo breve. Per contro, una sequenza stretta e un passo lungo denotano un animale che cammina in modo efficiente e si muove rapidamente. Una sequenza media sia dal punto di vista della larghezza sia da quello del passo può essere indizio o di un animale che cammina in modo non troppo efficiente o di un animale che cammina in modo efficiente ma si muove a una velocità relativamente ridotta.

Per le impronte di quadrupedi un'ulteriore misura importante è quella presa dal punto intermedio fra due impronte consecutive di un piede posteriore al punto intermedio fra due impronte analoghe di un piede anteriore. Si tratta in pratica della distanza fra l'articolazione dell'anca e quella della spalla dell'animale vivo, e corrisponde approssimativamente alla lunghezza del tronco dell'animale (esclusi il capo, il collo e la coda). Poiché il bordo delle impronte e dei calchi può sfumare uniformemente nella matrice circostante, spesso è difficile determinare i limiti esatti. Di conseguenza misure imprecise delle orme sono sin troppo comuni, e il ricercatore sul campo deve stare ben in guardia. Un altro rischio è dovuto al fatto che in alcuni casi l'animale con il suo peso ha lasciato l'impronta non solo sullo strato superficiale del sedimento ma anche su strati sottostanti. Bisogna quindi stare molto attenti a distinguere fra un'orma vera e propria e un'«orma fantasma». Il ricercatore dovrebbe esaminare non solo l'impronta ma anche un suo calco. Se ciò che è stato rinvenuto è proprio un calco, dovrebbe esaminare uno stampo ricavato dal calco stesso (che, naturalmente, ricrea l'impronta mancante). Se è una vera orma, allora bisognerebbe farne un calco. I calchi ottenuti sul campo offrono una preziosa documentazione per lo studio successivo.

Un altro aiuto all'analisi delle impronte è la misura del grado di sovrapposizione fra le orme dei piedi anteriori e di quelli posteriori. Ove si eccettuino certi animali con il corpo flessibile, come le salamandre, l'entità e la costanza della sovrapposizione danno una seconda misura della distanza tra l'anca e la spalla del quadrupede. Infine, un'altra misura utile è quella che definisce la posizione delle orme rispetto alla linea mediana della sequenza. Questo dato può fornire informazioni sull'andatura di un quadrupede. Se, per esempio, si tratta di un'andatura scomposta, la linea mediana della sequenza non corrisponde esattamente al punto medio tra orme appaiate.

Ricostruzione di animali sconosciuti

I rischi che si corrono quando si tenta di ricostruire un animale noto soltanto dalle sue orme fossili sono ben dimostrati da un esempio dell'Ottocento. L'animale è il *Chirotherium*, dal greco *kheir*, mano, e *therion*, animale selvatico. Le sue impronte furono scoperte per la prima volta nel 1834 in Turingia, nella Germania centrale, in un'arenaria rossa che risale al Triassico (da 225 a 195 milioni di anni fa). Le orme del piede posteriore hanno più o meno le dimensioni della mano di un uomo; quattro dita sono dirette in avanti, ma il quinto dito è disposto ad angolo rispetto agli altri quattro, più o meno come il pollice della mano umana. Alcune orme del piede posteriore erano associate con impronte più piccole ma molto simili, che furono scambiate per orme del piede anteriore.

Anche se gli animali terrestri prevalenti nel Triassico erano rettili, le orme fossili rinvenute in Turingia furono attribuite da principio a una qualche specie di scimmia antropomorfa o a un orso. Coloro i quali erano giunti a questa conclusione vi furono tratti da un'interpretazione errata: identificarono infatti il dito disposto ad angolo con un pollice, anche se esso si trovava all'esterno dell'orma. Alcune ricostruzioni raffigurarono addirittura il *Chirotherium* come un anfibio simile a un rospo che camminava a zampe incrociate; questo almeno serviva a mettere il «pollice» in una posizione normale. In una relazione presentata nel 1874 alla British Association for the Advancement of Science, lo zoologo L. C. Miall dimostrò che le ricostruzioni a zampe incrociate erano false e avanzò invece l'ipotesi che si trattasse di impronte di dinosauri. Che tali impronte fossero quanto meno di un rettile risultò evidente quando fu riconosciuto che la sequenza delle orme del *Chirotherium* era stretta. Le sequenze degli anfibi sono invariabilmente larghe.

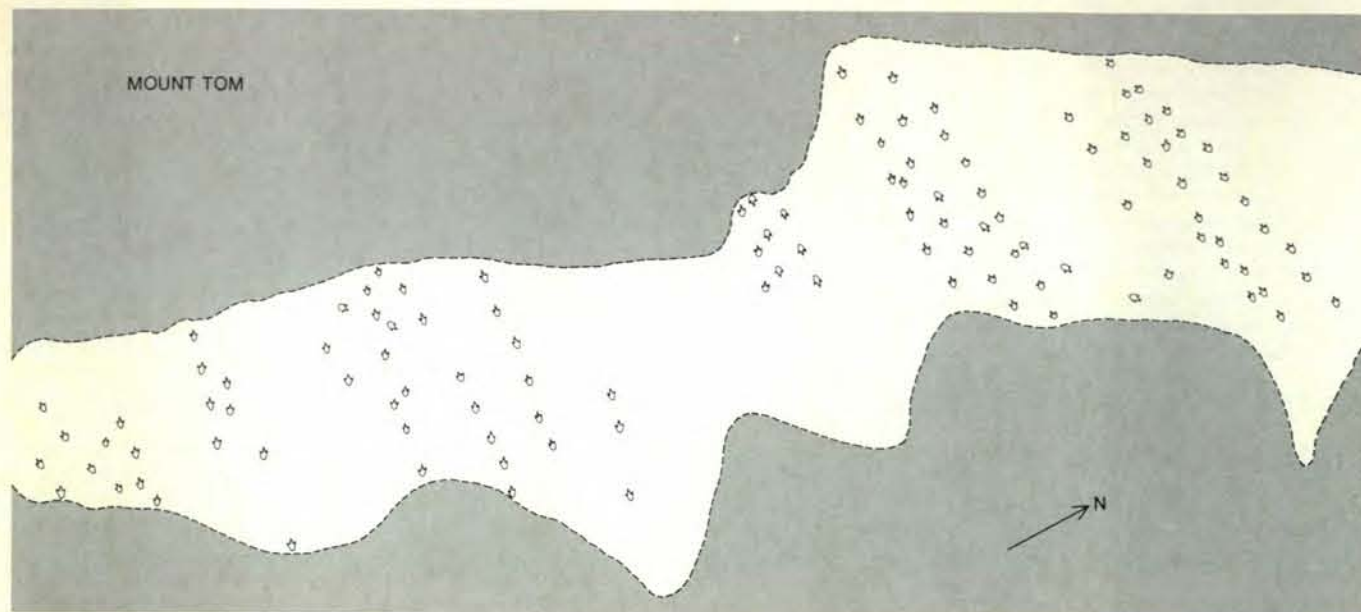
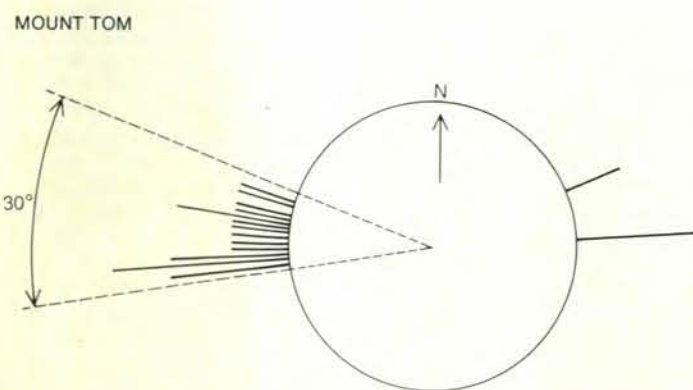
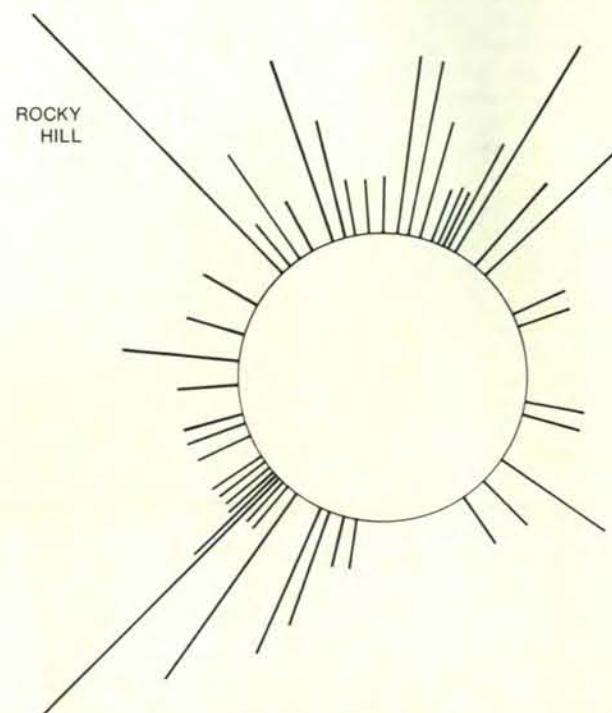
La prova più convincente che le impronte erano state lasciate da un rettile fu il fatto che le orme di *Chirotherium* in migliore stato di conservazione presentavano segni di squame. Mostravano anche l'esistenza di forti artigli e facevano così pensare che il rettile fosse carnivoro. L'analisi delle impronte induceva a ritenere che il corpo dell'animale fosse lungo circa un metro e che il rettile, dato che l'andatura era prevalentemente bipede, si mantenesse in equilibrio in posizione eretta grazie a una coda di pari lunghezza. Tenendo conto del capo e del collo dell'animale, si poteva attribuire al *Chirotherium* una lunghezza di circa due metri e mezzo dall'apice del muso alla punta della coda.

Oggi si sa che l'animale che lasciò quelle impronte era uno dei rettili pseudosuchi, un importante gruppo primitivo antenato tanto dei dinosauri quanto dei coccodrilli. Sebbene oggi quasi due dozzine di specie di *Chirotherium* vengano distinte in base alla forma dell'orma, si conoscono soltanto pochi scheletri fossili di tipo corrispondente. Uno di essi, il *Ticinosuchus*, scoperto meno di vent'anni fa in Svizzera in una formazione del Triassico medio, è uno dei principali candidati all'attribuzione delle impronte che sconcertarono i paleontologi dell'Ottocento.

Proprio come la struttura delle squame in certe orme di *Chirotherium* aiutò a stabilire l'identità dell'animale come rettile, così le orme di altri animali possono rivelare particolari di parti molli che solo raramente si fossilizzano. Un esempio è costituito dalle impronte di adrosauro rinvenute in formazioni del Cretaceo superiore nella provincia canadese di Alberta. Gli adrosauri erano dinosauri erbivori bipedi di grandi dimensioni. Le orme della provincia di Alberta indicano che, al pari degli elefanti del giorno d'oggi, questi animali avevano sviluppato dei cuscinetti carnosi sulla pianta e sulle dita dei piedi posteriori per ammortizzare il proprio rispettabile peso.

Impronte e comportamento

Le orme fossili hanno dimostrato di essere in grado sia di dare informazioni sia di correggere convinzioni errate in

merito al comportamento individuale e di gruppo degli animali che le lasciarono. I dinosauri del gruppo dei sauropodi, per esempio, descritti per la prima volta nel 1842 dallo zoologo inglese Richard Owen, comprendono i più grandi animali di terraferma che siano mai vissuti. Per più di un secolo si era dato per scontato che questi erbivori quadrupedi tozzi e pesanti, con il loro lungo collo e con la loro lunga coda, avessero condotto abitualmente vita acquatica traendo vantaggi dalla spinta idrostatica per sostenere nell'acqua il proprio enorme peso. Ci si chiedeva anzi seriamente se avessero potuto muoversi sulla terraferma. Poi Ronald T. Bird, dell'American Museum of Natural History, scoprì impronte di sauropodi in due formazioni texane del Cretaceo: a Paluxy Creek nel 1944 e a West Verde Creek nel 1954. I suoi reperti risolsero la questione in maniera decisiva. Una sequenza continua di orme di brontosauro rinvenuta a Paluxy Creek e ora esposta all'American Museum è particolarmente significativa e anche drammatica. Si tratta di una sequenza stretta, con

L'orientamento di queste sequenze continue di impronte rinvenute nella valle del fiume Connecticut nei pressi di Holyoke, nel Massachusetts, fa pensare all'esistenza di un comportamento sociale fra i dinosauri. L'analisi delle impronte fatta da John H. Ostrom della Yale University rivela che almeno 15 dinosauri carnivori del genere *Eubrontes* hanno attraversato la sabbia bagnata procedendo da est verso ovest. Il fatto che tutti gli animali si muovessero in un'unica direzione comune ha portato a ritenere che stessero cacciando in branco.

Diagrammi che mettono a confronto l'orientamento non casuale delle impronte di *Eubrontes*, a sinistra, con l'orientamento molto più casuale delle oltre 1000 orme analoghe di dinosauro rinvenute a Rocky Hill, un altro sito della valle del fiume Connecticut, a destra. Non tutte le impronte osservate nell'uno o nell'altro sito sono presentate nell'illustrazione. Nel diagramma di sinistra, comunque, il 70 per cento di tutte le sequenze continue si trovano in un arco di 30 gradi. L'evidente distribuzione casuale che si osserva nel diagramma di destra fa pensare che le sequenze di Rocky Hill siano state impresse in un periodo di tempo più lungo, cosa che non esclude per altro un'attività gregaria.

orme enormi impresse profondamente in quello che sarebbe potuto essere soltanto un piano fangoso, parzialmente o interamente esposto all'aria. I passi dell'animale, lunghi due metri buoni, sono inequivocabilmente quelli di un animale terrestre che camminava in maniera efficiente. Questo è tutto per quel che riguarda il presunto habitat acquatico abituale. Il dramma è offerto da qualcos'altro. Sovrapposte alle impronte del brontosauro vi sono le orme di un dinosauro bipede carnivoro di grandi dimensioni - forse un *Allosaurus* - che procedeva nella stessa direzione. Sembra probabile che il carnivoro stesse inseguendo l'erbivoro.

L'altro sito di Bird, West Verde Creek, presenta insieme a orme di sauropodo un solco profondo e continuo. Evidentemente il solco fu prodotto dalla coda del dinosauro trascinata per terra, segno che il piano fangoso era completamente allo scoperto e ulteriore conferma del fatto che i sauropodi potevano camminare senza bisogno di essere sostenuti dall'acqua. Ciò non significa peraltro che questi giganteschi dinosauri non procedessero mai a nuoto. Un altro reperto di Bird è significativo in proposito. Si tratta di una particolare sequenza di impronte di sauropodo che consiste unicamente di orme di piedi anteriori, ove si eccettui un'unica orma di un piede posteriore sinistro. A quanto pare la parte posteriore e la coda galleggiavano in acque abbastanza basse da permettere all'animale di camminare sul fondo con i piedi anteriori, proprio come fa oggi l'ippopotamo. Poi, nel girare a sinistra, il sauropodo appoggiò sul fondo un solo piede posteriore, lasciando l'impronta supplementare.

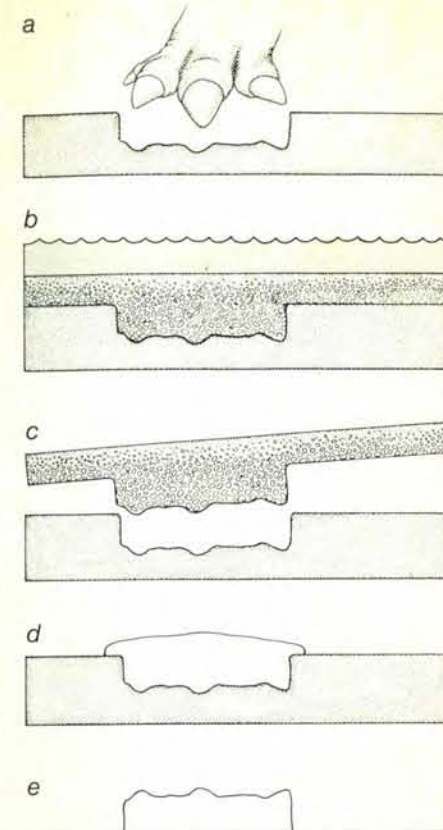
Nelle raffigurazioni tradizionali, questi giganteschi erbivori sono stati sempre presentati immersi nell'acqua nell'atto di nutrirsi con aria soddisfatta di piante acquatiche mentre i loro nemici naturali, i dinosauri carnivori, se ne stanno frustrati sulla sponda, incapaci di inseguire la preda. Alcune impronte conservate nello State Dinosaur Park di Rocky Hill, nel Connecticut, fanno pensare a una scena diversa. Lì centinaia di orme, per lo più del genere predatore *Eubrontes*, si sono conservate in sedimenti lacustri del Giurassico inferiore. Alcune sono perfettamente chiare, altre sono più indistinte; alcune sono impresse profondamente, altre sono poco profonde. Tutte presentano le tre dita caratteristiche dei dinosauri carnivori bipedi. Una sequenza in particolare, scoperta da Walter P. Coombs, Jr., del Western New England College, fa luce sull'abilità natatoria dei dinosauri carnivori. Essa consiste di otto orme consecutive in cui la lunghezza di ogni passo è in media di oltre un metro. Il passo con il piede destro avanti è regolarmente una ventina di centimetri più lungo di quello con il piede sinistro, cosa che ha indotto Coombs a ritenere che i movimenti natatori avessero una sorta di ritmo galoppante. Alcune di queste orme presentano inoltre segni eccezionalmente chiari di artigli, e ciò esclude la possibilità che si tratti di orme fantasma, anche se la loro configurazione è tale da far ritenere che il peso del corpo non gravasse sui piedi. Secondo l'interpretazione di Coombs, l'animale stava nuotando in acque piuttosto basse, toccando il fondo con le punte delle dita. La sequenza delle impronte inoltre incomincia e finisce bruscamente, indicando che il movimento natatorio sollevava a intermittenza l'animale dal fondo. Coombs conclude che la tradizionale immagine del carnivoro frustrato in piedi ai bordi dell'acqua ha bisogno di una revisione.

Le testimonianze fossili relative alla possibilità o meno che i dinosauri siano vissuti in gruppi sociali sono state molto limitate fino a pochissimo tempo fa. Un esempio di vita sociale è dato da un cumulo di oltre venti scheletri rinvenuti a Bernissart, in Belgio, in una formazione del Cretaceo inferiore. Tutti gli animali erano grandi erbivori bipedi del genere *Iguanodon*. È stata avanzata l'ipotesi che fossero morti quando un intero branco cadde in un burrone, tacita testimonianza del fatto che gli iguanodonti viaggiavano in branchi. Recentemente David Norman, dell'Università di Oxford, ha dimostrato che gli animali non furono rinvenuti tutti allo stesso livello della formazione e che la formazione stessa è in realtà una serie di sedimenti. Le sue scoperte fanno sorgere seri dubbi sull'interpretazione relativa alla vita in branchi.

Altri dinosauri sociali

Un secondo esempio è costituito da un cumulo di scheletri di dinosauri rinvenuti nel 1947 a Ghost Ranch, nel New Mexico, in una formazione triassica. Tutti gli animali appartenevano allo stesso genere carnivoro, *Coelophysis*, di piccole dimensioni. Alcuni vedono anche in questo reperto una prova del fatto che quei carnivori cacciavano in branchi, mentre altri non sono del medesimo parere. A dire il vero, per la loro stessa natura fossile, non si può ritenere che resti di scheletri possano fornire una testimonianza diretta del comportamento di un animale.

Una fonte è costituita dai «nidi» fatti dai dinosauri. Le scoperte in Mongolia negli anni venti di nidi di dinosauri con tanto di uova non ancora dischiuse, disposte simmetricamente come quelle degli uccelli, fu il fatto sensazionale del momento in campo paleontologico. Uno di questi nidi, quello di un dinosauro quadrupede del genere *Protoceratops*, attestava il fatto che, a differenza di certi rettili moderni come le tartarughe, non è da escludere che questi animali cretacei custodissero le loro uova non ancora dischiuse. Accanto al nido c'era lo scheletro fossilizzato di un celurosauro predatore - un dinosauro carnivoro bipede di piccole dimensioni - che a quanto pare era andato incontro a morte violenta. Il suo scopritore attribuì il fossile a un nuovo genere, a cui diede appropriatamente il nome di *Oviraptor*.

Una seconda scoperta di un nido, fatta soltanto due anni fa nel Montana, continua la storia della socialità dei dinosauri. Nella formazione di Two Medicine, un

La conservazione delle orme richiede, una volta lasciata l'impronta *(a)*, che il sedimento intaccato sia riempito da un sedimento di struttura diversa *(b)* e che entrambi i sedimenti si induriscano. I due sedimenti si possono in seguito separare. Se si conserva soltanto il sedimento riempiente *(c)* si tratta di un calco naturale. Se invece si conserva soltanto il sedimento intaccato *(d)* si tratta di un'orma naturale, che può essere poi usata per fare dei calchi *(e)*.

deposito del Cretaceo, John R. Horner, della Montana State University, rinvenne gli scheletri di 18 piccoli di adrosauro. Si ritenne che i giovani animali avessero una o due settimane di vita. La loro associazione con un nido fa pensare a un comportamento sociale di carattere familiare, e comporta anche la possibilità che i genitori si prendessero cura dei piccoli quando questi nascevano.

Quale comportamento sociale si può dedurre dalle orme? Tanto per cominciare, la prima chiara testimonianza di una vita in branchi dei dinosauri è venuta dagli stessi due siti texani studiati da Bird. Una ventina delle impronte di sauropodo scoperte da Bird erano parallele fra loro, segno che gli animali procedevano in una direzione comune, e alcune si sovrapponevano. Le orme differivano inoltre dal punto di vista delle dimensioni, dimostrando che il branco comprendeva tanto animali giovani quanto animali maturi.

Alcuni anni dopo la scoperta di Bird, John H. Ostrom, della Yale University, scoprì a Lake Eanes (Texas), in una formazione geologica del Cretaceo inferiore, 25 serie di orme di dinosauro. Tutte erano state lasciate da erbivori bipedi, proba-

Due sequenze di impronte scoperte nel 1944 da Ronald T. Bird a Paluxy Creek, nel Texas, registrano un episodio drammatico del Mesozoico. La serie di orme più grandi è quella di un dinosauro erbivoro quadrupede, probabilmente un brontosauro. Le orme tridattili più piccole visibili a sinistra della sequenza principale furono lasciate da un dinosauro carnivoro bipede, forse del genere *Allosaurus*. Il carnivoro si muoveva nella stessa direzione dell'erbivoro, probabilmente al suo inseguimento.

bilmente del genere *Camptosaurus*. Tutte erano orientate nella stessa direzione, ossia verso nord-ovest. Ostrom ha trovato testimonianze che fanno pensare all'uso di cacciare in branchi anche fra i dinosauri carnivori conservati nei sedimenti triassici di Mount Tom nel Massachusetts. Qui, sulla superficie di un piano di stratificazione, si trovano almeno 19 impronte distinte di dinosauri del genere *Eubrontes*. Anche qui tutti gli animali procedevano nella stessa direzione. Evidentemente camminavano affiancati, poiché nessuna delle sequenze di impronte si sovrappone.

Così, la testimonianza offerta dalle orme rende chiaro che alcuni dinosauri erbivori e carnivori erano gregari. Questo è di per sé un comportamento sociale di tipo avanzato che non ha riscontro fra i rettili del giorno d'oggi. Sulla base delle impronte texane di Bird, Robert T. Bakker, della Johns Hopkins University, ha avanzato addirittura l'ipotesi che questi sauropodi viaggiassero in branchi organizzati con i giovani al centro e gli adulti disposti in modo da formare un perimetro protettivo. A questo punto sarebbe più che benvenuta qualche prova aggiuntiva dalla quale risultasse che la disposizione organizzata non era casuale.

Un ricco giacimento canadese

L'inventario delle orme fossili si è molto arricchito in questi ultimi anni grazie a lavori intensivi svolti sui sedimenti del medio Cretaceo del Peace River Canyon, nella British Columbia. Il canyon, lungo 28 chilometri, è noto come luogo ricco di impronte da quando un pioniere della paleontologia, Charles M. Sternberg, del Royal Ontario Museum, vi scoprì nel 1922 quattrocento orme di dinosauri. Sternberg rese di pubblico dominio le sue scoperte nel 1930, dando un nome a otto nuovi tipi di impronte, tutte lasciate da dinosauri e tutte, tranne una, tali da dimostrare una locomozione bipede.

Nel 1975 cominciò la costruzione di una grande diga a valle del sito delle scoperte di Sternberg. Questo fatto spinse il Natural History Museum della provincia di Alberta a intensificare gli sforzi per salvare almeno alcune delle altre orme fossili che si presumeva esistessero nelle parti non ancora esaminate del canyon, prima che questo venisse inevitabilmente inondato. Cinque stagioni di lavoro sotto la guida di Philip Currie hanno dato risultati spettacolari.

Sebbene fosse possibile studiare soltanto i sette chilometri inferiori del canyon, furono trovate più di 100 sequenze di impronte. Esse comprendevano in tutto 1700 orme di dinosauri caratteristiche di almeno 10 specie diverse, una serie di orme di uccello e una sola sequenza di orme di tartaruga. Più di 1000 di queste impronte sono state segnate su una mappa e quasi altrettante sono state misurate. È stato fatto il calco di circa 200 orme, e 90 sono state rimosse per essere studiate più a fondo e per essere alla fine esposte al pubblico.

Fra le splendide impronte di dinosauro c'è un'intera serie che permette di seguire la crescita - da giovane a adulto - di un adrosauro primitivo. Questa serie rivela come la forma dei piedi si modificasse a mano a mano che gli animali diventavano adulti e a mano a mano che il loro peso aumentava: una documentazione senza precedenti in icnologia. Pare che un branco di questi erbivori abbia attraversato i sedimenti cretacei, procedendo a volte in ordine sparso e a volte fianco a fianco. A giudicare dalla frequente sovrimpressione delle orme dei piccoli su quelle degli adulti, gli animali più giovani seguivano quelli più maturi. Una serie di impronte indica che 11 animali procedettero per un po' verso sud e poi di colpo presero ad andare verso est. (La conservazione di una sequenza singola di impronte di un adrosauro che procedeva nella direzione opposta, da sud a nord, indica che la svolta a sinistra degli altri non era stata imposta da qualche ostacolo naturale.) Un attento esame di 4 delle 11 sequenze ha rivelato addirittura che quando uno degli adrosauri sbandò da una parte, tre degli altri cambiarono direzione per evitare lo scontro.

Le orme di Peace River hanno portato nuovi argomenti relativi alla possibilità che dinosauri carnivori di piccole e medie dimensioni cacciassero in branchi. In una località, per esempio, sei sequenze continue di impronte di carnivori di dimensioni medie sono dirette tutte nella stessa direzione, sebbene due tracce si incrocino un paio di volte. Per contro, le impronte di carnivori di maggiori dimensioni - probabilmente tirannosauri - rinvenute nel canyon sono quelle di animali singoli o al massimo di coppie. Non esiste pertanto ancora nessuna prova, basata su orme rinvenute in qualsiasi località del mondo, che indichi che i dinosauri carnivori di maggiori dimensioni cacciassero in branchi.

Le orme possono pure indicare la velocità alla quale l'animale che le ha lasciate si stava muovendo, anche se bisogna stare molto attenti nelle interpretazioni. Un esempio significativo a questo proposito è costituito dalle impronte di dinosauri scoperte agli inizi degli anni sessanta in una cava inglese nei pressi di Herston, nel Dorset, dove si trovano i Purbeck Beds, una formazione geologica del Cretaceo superiore. Sul fondo della cava furono rinvenute due linee parallele di orme di un erbivoro bipede. L'interpretazione iniziale fu che si trattava delle orme dei piedi destro e sinistro di un unico animale. La lunghezza ridotta del passo completo fece ritenere che il dinosauro si muovesse molto pesantemente; vi fu chi avanzò addirittura l'ipotesi che, siccome poggiava chiaramente tutto il peso del corpo sulla rotondità del piede, l'animale stesse superando quella che allora era una pendenza ripida. Questo quadro si adattava all'immagine allora corrente dei dinosauri quali animali lenti e pigri. Quando però le ricerche si estesero, si scoprì che le impronte divergevano: erano state lasciate non da un solo dinosauro che procedeva lentamente in salita, ma da due dinosauri molto più veloci che forse correvano fianco a fianco.

Le testimonianze delle impronte di Peace River rendono chiaro che alcuni dinosauri erano in grado di muoversi molto rapidamente. Con una formula elaborata da R. McNeill Alexander, dell'Università di Leeds, è possibile calcolare la velocità del movimento dell'animale osservando la disposizione delle sue orme e la distanza fra esse. Le velocità più elevate calcolate in base alle tracce di Peace River sono quelle di carnivori di medie dimensioni: 16,5 chilometri all'ora. Altri carnivori si muovevano a velocità più ridotte: tra 6 e 8,5 chilometri orari. Gli erbivori si muovevano più lentamente. La loro velocità massima era costituita da quegli stessi 6 chilometri orari espressi dall'andatura dei carnivori più lenti.

In complesso, quindi, i recenti studi sulle orme dimostrano che il quadro tradizionale dei dinosauri quali animali lenti e perfino goffi e ottusi è ben lontano dal vero. Sembra probabile che, quando si saranno fatte analisi altrettanto approfondite delle impronte fossili di altri vertebrati terrestri, ad esempio di anfibi, di rettili d'altre specie e di mammiferi del Cenozoico inferiore, saranno disponibili, sul loro comportamento, nuove informazioni che risulteranno altrettanto in disaccordo con le conclusioni passate. Siamo forse agli inizi di una rinascita dello studio delle orme fossili? Sembra comunque facile prevedere che nei prossimi decenni questo campo di ricerca sarà uno dei più stimolanti della paleontologia.

Spettroscopia NMR delle cellule viventi

Tradizionalmente le reazioni chimiche del metabolismo sono studiate a una a una in vitro; con la spettroscopia a risonanza magnetica nucleare (NMR) le stesse reazioni possono essere seguite in vivo

di R. G. Shulman

Verso la metà del XIX secolo Louis Pasteur dimostrò che le cellule viventi sono in grado di produrre specifiche reazioni chimiche: dimostrò che il lievito, che fermenta i carboidrati trasformandoli in sostanze quali l'anidride carbonica e l'alcool etilico, è costituito di cellule viventi. Alla fine del secolo, i fratelli Eduard e Hans Buchner macinarono delle cellule di lievito e dimostrarono che i carboidrati, che venivano aggiunti alla soluzione così ottenuta, erano convertiti negli stessi prodotti finali. In questo modo essi stabilirono che alcune sostanze contenute nelle cellule di lievito catalizzano specifiche reazioni chimiche. La scoperta aprì nuove strade al progresso della biochimica. Durante tutto il secolo attuale i biochimici hanno tentato di spiegare il modo in cui alcune sostanze estratte dalle cellule e purificate potrebbero controllare le reazioni chimiche di cui la cellula è capace. Numerosi studiosi hanno identificato diverse reazioni, chiarendo alcune delle vie metaboliche nell'insieme complesso degli eventi che costituiscono il metabolismo intracellulare. Nonostante alcuni aspetti del metabolismo rimangano inesplorati, la scoperta di vie metaboliche, che si incrociano, si congiungono e si configurano in cicli permette di avvicinarci sempre più alla comprensione della complessità della cellula vivente.

Nella misura in cui persistono obiezioni sulla validità della biochimica moderna, esse riguardano sempre la riduzione dei processi vitali a sequenze di reazioni chimiche. «Queste reazioni possono aver luogo in vitro - si sente dire - ma le cose vanno realmente così nelle cellule viventi? E che cosa accade negli organismi pluricellulari?» Descriverò qui come una nuova tecnica cominci a rispondere a queste domande rilevando le reazioni chimiche man mano che esse avvengono nelle cellule, nei tessuti e negli organismi, compreso quello umano. La tecnica è la spettroscopia a risonanza magnetica nucleare (NMR, *nuclear magnetic resonance*) e si basa sul fatto che i nuclei degli atomi con un numero dispari di nucleoni (protoni e neutroni) hanno un magnetismo intrinseco che fa di ciascun nucleo un dipolo magnetico; in sostanza una sbarretta magnetica. Tali nuclei includono il protone (H-1), che è il nucleo del 99,98 per cento di tutti gli atomi di idrogeno che si trovano in natura, il nucleo del carbonio 13 (C-13), che è il nucleo dell'1,1 per cento di tutti gli atomi di carbonio, e il nucleo del fosforo 31 (P-31), che è il nucleo di tutti gli atomi di fosforo.

Nella spettroscopia NMR vengono applicati alle cellule, ai tessuti o alle parti di un organismo vivente da esaminare due campi. Il primo è un intenso campo magnetico che induce i dipoli nucleari (cioè i nuclei dell'idrogeno, del carbonio 13 e del fosforo 31 nel campione da esaminare) a disporsi in modo tale per cui il dipolo di ciascun nucleo sia allineato (ossia orientato) o nel verso del campo o nel verso opposto a esso. L'orientazione nel verso del campo è uno stato in cui il nucleo immagazzina meno energia rispetto a quando esso è orientato nel verso opposto al campo.

Si applica, quindi, il secondo campo. Esso consiste di radiazioni elettromagnetiche appartenenti alla regione delle frequenze radio dello spettro. Per ogni data intensità del campo magnetico, in cui è posto il campione da esaminare, c'è una data frequenza di radiazioni elettromagnetiche, per la quale ciascun fotone, o quanto di radiazione, trasporterà l'esatta quantità di energia necessaria per permettere che l'orientazione di un certo tipo di nucleo si inverta. Quindi, se l'intensità del campo magnetico è tenuta costante e si fa variare la frequenza delle onde radio (o, al contrario, la frequenza delle onde radio è tenuta costante e si fa variare l'intensità del campo magnetico), c'è un momento in cui i nuclei risuonano: assorbono i fotoni delle onde radio. Per esempio, in un campo magnetico con una intensità di 84 000 gauss un nucleo di idrogeno 1 risuona alla frequenza di circa 360 megahertz (360 milioni di cicli al secondo), un nucleo di fosforo 31 a circa 146 megahertz e un nucleo di carbonio 13 a circa 90 megahertz.

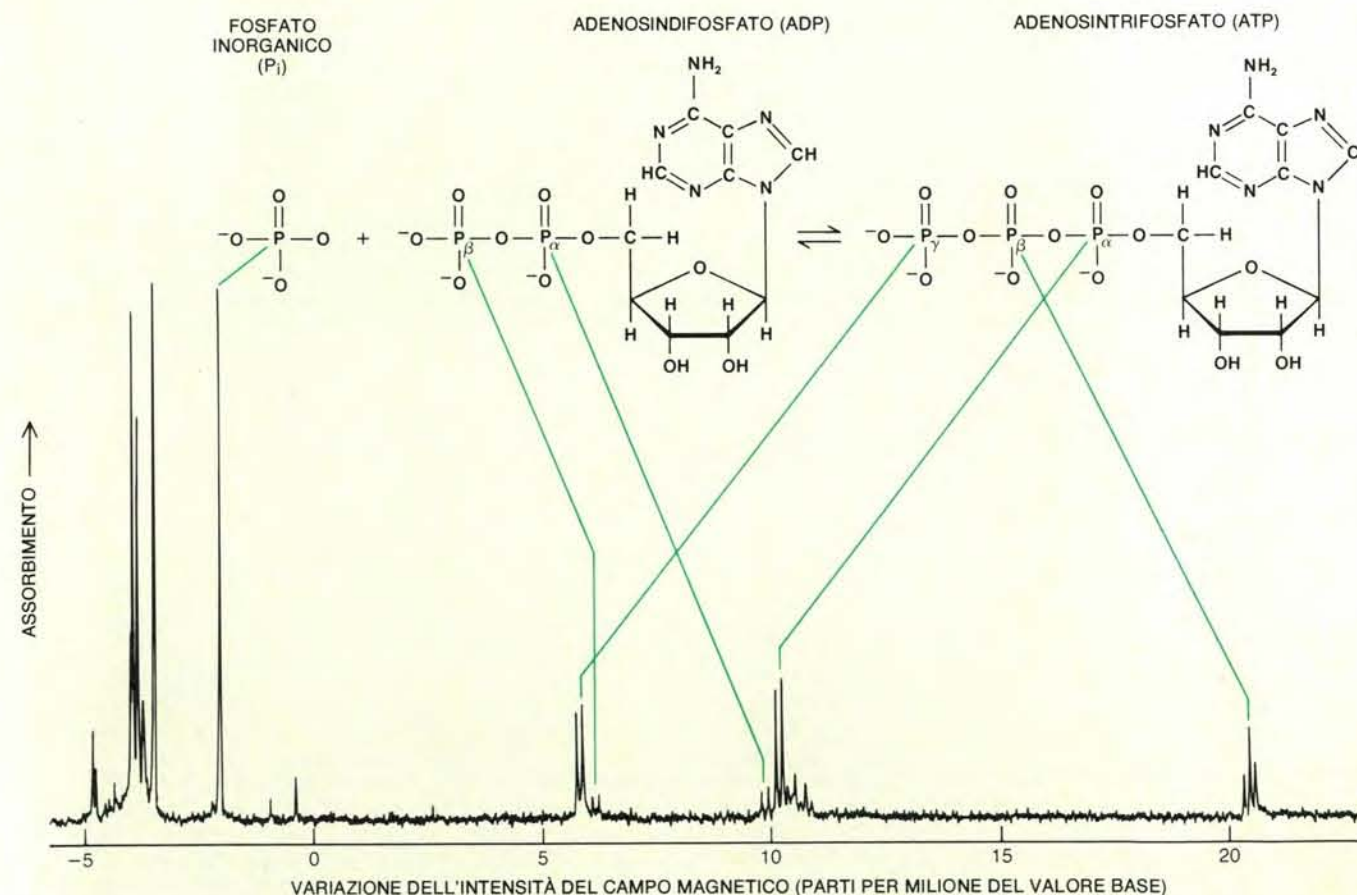
Nelle cellule, nei tessuti e negli organismi viventi la frequenza precisa di risonanza dipende da un'ulteriore circostanza: l'ambiente chimico in cui è situato il nucleo. La frequenza ha uno «spostamento chimico», cioè differisce da quella che ci si aspetterebbe se il nucleo fosse isolato, perché nelle cellule, nei tessuti e negli organismi viventi i nuclei fanno parte di atomi e gli atomi di molecole. Si consideri uno spettro NMR che mostri le risonanze (in forma di picchi, o di righe di assorbimento) di un campione di acido glutammico, un amminoacido. Benché il carbonio 13 abbia una abbondanza in natura di solo l'1,1 per cento, appaiono chiaramente i picchi dovuti al C-13 delle cinque posizioni del carbonio nella molecola di glutammato e, se si ottiene lo spettro variando l'intensità del campo magnetico, essi si estendono in un intervallo di circa 200 parti per milione (0,02 per cento) dell'intensità del campo. Nella molecola di glutammato due degli atomi di carbonio sono legati ciascuno a due atomi di ossigeno (COO^-); i loro picchi sono vicini. Tuttavia questi picchi sono separati abbastanza dal picco di un atomo di carbonio che è legato a un atomo di idrogeno e dai picchi di due atomi di carbonio, ciascuno dei quali è legato a due atomi di idrogeno.

In uno spettro NMR un dato picco può essere assegnato a un atomo di carbonio in una particolare posizione di una data molecola tramite varie tecniche. La tecnica più diretta è quella di misurare lo spettro dopo che una particolare posizione è stata marcata con carbonio 13, così che il numero degli atomi di C-13 in quella posizione in un campione della sostanza sia di gran lunga superiore a quello che si potrebbe prevedere in base alla abbondanza naturale del C-13. Il picco di assorbimento del carbonio in quella posizione sarà corrispondentemente accentuato. La tecnica presenta, nel caso delle reazioni chimiche del metabolismo, il vantaggio di poter seguire il C-13 introdotto come tracciante mentre passa attraverso varie posizioni in varie molecole.

Il nucleo del fosforo 31 presenta diversi vantaggi. Poiché il nucleo del P-31 è il 100 per cento di tutti i nuclei di fosforo, non sono necessarie tecniche di marcatura. Inoltre, il nucleo del P-31 dà forti segnali NMR per alcuni composti importanti. Quando si applica la spettroscopia NMR a una sospensione del batterio *Escherichia coli*, si rilevano tipicamente picchi dai tre atomi di fosforo dell'adenosintrifosfato (ATP), una molecola in cui la cellula immagazzina energia chimica. Si rileva un picco dall'atomo di fosforo del fosfato inorganico (P_i), che si stacca dall'ATP quando viene scisso il legame ad alto contenuto energetico. Si rilevano picchi dai due atomi di fosforo del nicotinammideadenindinucleotide, una molecola che partecipa alle reazioni di ossidazione intracellulare. A volte si rilevano picchi dall'adenosindifosfato (ADP), che ha un atomo di fosforo in meno rispetto all'ATP; dal fosfoenolpiruvato, dal quale la cellula può staccare il P_i necessario per convertire l'ADP in ATP, e da catene di polifosfati.

Quando i batteri vivono basandosi su una provvista di glucosio, essi utilizzano la sequenza di reazioni metaboliche nota come glicolisi o via glicolitica. Anche le molecole che si presentano lungo la sequenza, in particolare quella del fruttosio-1,6-difosfato (FDP), possono essere rilevate. Tali molecole hanno picchi NMR molto netti, poiché ciascuna molecola è piuttosto piccola e piccole molecole in soluzione vengono demolite rapidamente (nell'ordine di 10^9 al secondo). La demolizione rapida significa che i legami chimici della molecola coprono tutti gli angoli possibili rispetto all'ambiente della molecola. In effetti, la demolizione corrisponde all'incirca alla somma degli angoli di legame rispetto all'ambiente. Una molecola di grandi dimensioni viene, invece, demolita lentamente e una molecola che fa parte di un solido non viene demolita affatto. I picchi di assorbimento di un gran numero di tali molecole sono allargati grazie al contributo che deriva dai nuclei delle molecole orientate ad angoli diversi rispetto al loro ambiente. In tali sostanze persino una grande concentrazione di nuclei di fosforo 31 può quindi non essere rilevata negli spettri NMR perché i picchi di assorbimento sono semplicemente troppo larghi.

Tra i picchi di fosforo 31 osservabili, quello dovuto al fosfato inorganico è stato il più semplice da interpretare e tra i più ricchi di informazioni. Il fosfato inorganico si trova all'interno delle cellule legato a uno o a due protoni. A un *p*H intracellulare di 6,7 le popolazioni della forma legata a un solo protone e a due protoni sono uguali. A un *p*H più alto (cioè in ambiente più alcalino), prevale la forma legata a un protone, e a un *p*H più basso (cioè in ambiente più acido), prevale la forma legata a due protoni. Ciascuna forma fornisce una risposta chimica tipica, ma il picco del fosfato inorganico in uno spettro NMR ottenuto da un campione di cellule si troverà in una posizione intermedia che dipende dal grado medio di protonazione di tutto il P_i delle cellule in esame.

La posizione del picco di P_i dà quindi una misura accurata del grado di protonazione del P_i, che a sua volta dà una misura accurata del *p*H intracellulare. Alcuni anni fa, lavorando con me presso i Bell Laboratories, Gil Navon e Kamil Ugurbil dimostrarono seguendo questa via che il

I picchi di uno spettro NMR rappresentano l'assorbimento di onde radio da parte dei nuclei di atomi situati in specifiche posizioni all'interno di molecole poste in un campo magnetico. La parte superiore dell'illustrazione mostra la reazione chimica in cui il fosfato inorganico (P_i) si unisce all'adenosindifosfato (ADP) e forma l'adenosintrifosfato (ATP), una molecola in cui le cellule viventi immagazzinano energia chimica. La parte inferiore dell'illustrazione mostra uno spettro NMR del fosforo 31 (P-31) ottenuto dal contenuto di cellule tumorali di topo. Quasi tutti i picchi di assorbimento dello spettro sono dovuti ai nuclei degli atomi di P-31 delle molecole riportate in alto. Il picco dovuto al P-31 in posizione gamma dell'ATP e il picco dovuto al P-31 in beta dell'ADP sono vicini perché questi atomi di fosforo hanno un ambiente chimico simile. Il picco dovuto al P-31 in beta dell'ATP è più isolato in quanto il suo ambiente chimico è unico essendo fiancheggiato da due gruppi fosfato.

*p*H all'interno di *Escherichia coli* è tenuto abbastanza costantemente intorno a 7,4, quando il batterio ha una fonte di energia come il glucosio. Negli stessi spettri NMR può essere seguita la conversione del glucosio in FDP, così come può essere seguito il conseguente aumento della concentrazione di ATP man mano che la cellula immagazzina l'energia che ha acquisito.

Il biochimico inglese Peter Mitchell ha ipotizzato che nella cellula la conversione di ATP in ADP e fosfato inorganico sia accoppiata al passaggio di protoni attraverso la membrana cellulare. In particolare, la reazione che libera energia chimica dalla conversione dell'ATP in ADP e P_i serve per espellere protoni dalla cellula, creando un gradiente di protoni che accumula sia energia fisica sia chimica. Viceversa, il rientro di protoni nella cellula fornisce l'energia richiesta per la reazione che combina ADP e P_i e forma ATP. Un cambiamento della concentrazione di protoni all'interno della cellula corrisponde a un cambiamento del *p*H intracellulare; quindi la risposta chimica del picco di P_i negli spettri NMR fornisce un banco di prova per l'ipotesi di Mitchell.

Si comincia la prova somministrando glucosio a batteri *E. coli* che hanno esaurito le loro riserve nutritive. In tali cellule il *p*H sarà sceso a valori considerevolmente inferiori a 7,0. Quando le cellule cominciano a metabolizzare la loro nuova provvista di glucosio, il picco dovuto al fosfato inorganico appare sdoppiarsi in due picchi. Uno di questi, corrispondente a un *p*H di 7,4, deriva dal P_i presente all'interno delle cellule. L'altro, che corrisponde a un *p*H più basso (cioè a una maggiore concentrazione di protoni), deriva dal P_i del mezzo in cui le cellule sono in sospensione. I picchi confermano che comincia a stabilirsi un gradiente di protoni. Contemporaneamente i picchi dovuti all'ATP stanno aumentando.

La reazione che scinde l'ATP in ADP e fosfato inorganico è catalizzata dall'ATP-asi, un enzima che media anche il trasporto di protoni attraverso la membrana cellulare. Se a una sospensione di *E. coli* si aggiunge un inibitore dell'ATP-asi insieme a nuovo glucosio, si osserva che il glucosio è anche in questo caso consumato. Presto compaiono FDP e quindi ATP. Persino quando si sono formate quantità apprezzabili di ATP, le osservazioni del picco di P_i rivelano che non compare un gradiente di protoni. Inoltre, la provvista di glucosio dura molto più a lungo rispetto alla prova precedente. Questo dimostra che una frazione sensibile del consumo di ATP intracellulare viene utilizzata per mantenere il gradiente di protoni.

L'importanza degli spettri NMR nello studio del metabolismo è particolarmente ben illustrata da una serie di esperimenti eseguiti sul muscolo da George K. Radda, David G. Gadian e colleghi presso l'Università di Oxford. Nel 1974 questi ricercatori dimostrarono che gli spettri NMR potevano essere utilizzati per osservare in vivo il metabolismo nei tessuti muscolari di rana. I loro spettri rilevarono picchi di fosforo 31 dovuti al fosfato inorganico, all'ATP e alla fosfocreatina (PCr), una molecola che le cellule muscolari utilizzano per convertire l'ADP in ATP. Quando il muscolo fu privato di una fonte di energia esterna, gli spettri mostrarono una perdita di PCr e un livello costante di ATP, mantenuto tramite la riserva di PCr. Solo quando la PCr fu esaurita, il livello di ATP scese, mentre rimanevano l'ADP e il P_i. Le posizioni dei picchi di ATP rivelarono che essenzialmente tutto l'ATP delle cellule era legato a ioni metallici carichi positivamente (probabilmente ioni magnesio).

Questi studi preliminari (completati da esperimenti simili eseguiti da Michael Barany, C. Tyler Burt e Thomas Glonek del Medical Center dell'Università dell'Illinois) riuscirono così a valutare gli stessi parametri che ricerche precedenti avevano valutato tramite analisi chimiche di numerosi campioni di muscolo, prelevati a vari intervalli dopo che il tessuto era stato nutrito. Seguirono ulteriori studi. Poiché la funzione del muscolo è quella di compiere un lavoro, Gadian, con Douglas Wilkie e Joan Dawson della University College London School of Medicine, cercò di utilizzare i picchi del fosforo 31 negli spettri NMR per rilevare le correlazioni tra i metaboliti e l'affaticamento del muscolo. Essi sospesero un muscolo di rana in un recipiente molto simile a una provetta e lo posero in uno spettrometro NMR.

In alcuni esperimenti la soluzione che circondava il muscolo era ossigenata, in altri no. Un elettrodo stimolava il muscolo causandone la contrazione. Un estensimetro misurava la forza sviluppata dal muscolo. Gadian e i suoi colleghi conclusero che una diminuzione della forza è ben correlata con un aumento della concentrazione di ADP. Inoltre esso è proporzionale al tasso a cui l'ATP è consumato. D'altronde, non dipende dalla concentrazione di ATP. Gli sperimentatori potevano seguire il tasso a cui l'ATP era consumato misurando la diminuzione del livello di fosfocreatina, la fonte di ATP del muscolo, e misurando contemporaneamente il livello di fosfato inorganico, che deriva dal consumo di ATP.

Negli ultimi due anni i progressi della spettroscopia NMR e la costruzione di nuovi spettrometri hanno permesso che gli studi del metabolismo del muscolo si estendessero anche ai muscoli dell'uomo. I precedenti spettrometri, costruiti per studi chimici, erano stati progettati per accogliere i campioni da esaminare in un tubo cilindrico. All'interno dello spettrometro il tubo con il suo contenuto era circondato da una bobina che emetteva radiazioni a frequenza radio e il tubo e la bobina erano immersi in un campo magnetico omogeneo statico. Chiaramente l'esame degli arti interi del corpo umano richiederebbe una strumentazione differente.

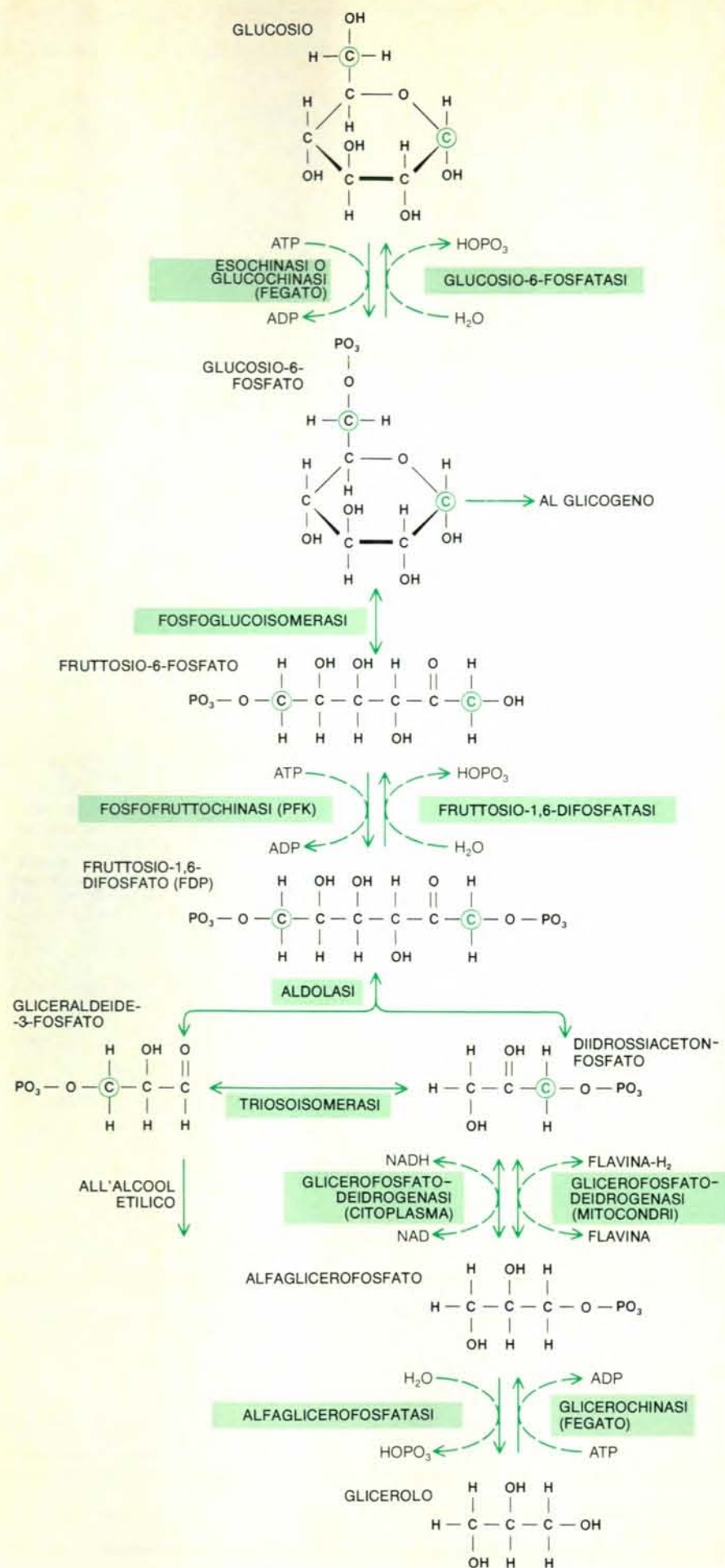
Nei nuovi spettrometri una bobina piatta è posta sulla superficie del campione da esaminare. La bobina piatta non eccita con onde radio l'intero campione, ma irradia solo la parte situata nelle sue immediate vicinanze. Questo è un vantaggio: lo sperimentatore può seguire le reazioni chimiche in una specifica zona del tessuto. Per circoscrivere ulteriormente la regione, il campo magnetico dei nuovi spettrometri può essere regolato in modo da essere omogeneo solo in un volume di uno o due centimetri cubi. Da questo volume sono rilevati netti picchi NMR. Nelle restanti parti del volume sotto l'influenza del magnete il campo magnetico è piuttosto disomogeneo; in questo modo, se le onde radio vi sono assorbite, l'assorbimento è diffuso su un ampio intervallo di frequenze o di intensità del campo magnetico e si perde nel rumore di fondo dello spettro.

Grazie alla superficie piatta della bobina e al campo magnetico focalizzato, Radda e Gadian, in collaborazione con i ricercatori degli Oxford Research Systems, la società che costruisce i nuovi spettrometri, sono stati in grado di misurare i picchi di assorbimento del fosforo 31 dal fegato di ratti vivi, escludendo i picchi dei tessuti muscolari circostanti. I loro spettri rappresentavano soltanto il fegato perché non mostravano picchi derivanti dalla fosfocreatina, la quale è assente nel fegato, ma è presente nel muscolo in alta concentrazione. Più recentemente Radda, lavorando con un gruppo di ricercatori clinici guidati da Brian D. Ross di Oxford, ha introdotto l'applicazione della spettroscopia NMR nella diagnosi delle malattie umane. In particolare, i ricercatori hanno diagnosticato la sindrome di McArdle in un paziente di 51 anni. La sindrome di McArdle è una malattia genetica in cui l'enzima fosforilasi è inattivo. La fosforilasi partecipa alla trasformazione del glicogeno in glucosio su richiesta delle cellule muscolari. Chi è affetto da questa sindrome presenta ipostenia dopo il più leggero esercizio muscolare.

Per giungere alla diagnosi, al paziente

La magnetizzazione intrinseca di alcuni nuclei (inclusi i nuclei dell'idrogeno 1, del carbonio 13 e del fosforo 31) rende possibile la spettroscopia NMR, in quanto ciascuno di tali nuclei è l'equivalente di una sbarra magnetica. In un campo magnetico di notevole intensità i nuclei si orientano o nel verso del campo o nel verso opposto a esso *(a)*. Se i nuclei vengono poi irradiati con onde radio, essi assorbiranno i quanti di radiazione che danno loro l'energia per invertire l'orientazione *(b)*. Il tipo di nucleo e il suo ambiente chimico determinano le combinazioni della frequenza radio e dell'intensità del campo magnetico a cui gli assorbimenti appaiono come picchi (o «righe») nello spettro NMR.

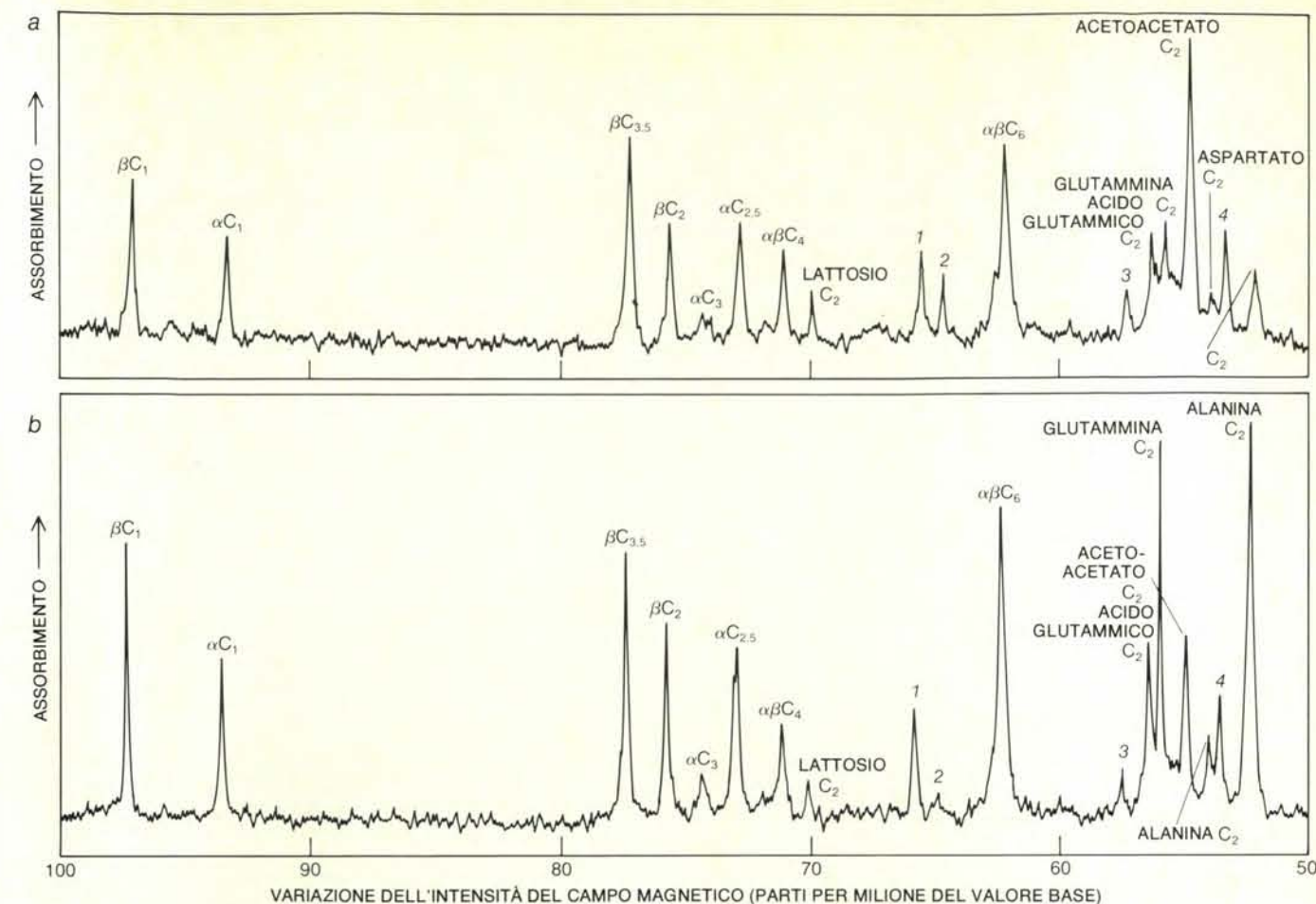
Gli effetti del consumo di glucosio sui composti contenenti fosforo nel batterio *Escherichia coli* sono rappresentati in questi spettri NMR del fosforo 31, che sono stati rilevati prima della somministrazione di glucosio alle cellule *(a)* e, quindi, da quattro a sei minuti dopo la somministrazione *(b)*. Nel secondo spettro il picco relativo al fosfato inorganico appare sdoppiato. Dalle posizioni dei due picchi risultanti si può dedurre che l'interno delle cellule è diventato più alcalino rispetto al mezzo in cui le cellule sono in sospensione. I picchi rimanenti sono dovuti a sostanze intracellulari come ADP, ATP e fruttosio-1,6-difosfato (FDP).

Sono qui schematizzati i meccanismi intracellulari ritenuti alla base dei cambiamenti degli spettri NMR mostrati nell'illustrazione in alto in questa pagina. Il metabolismo del glucosio comporta un aumento del FDP e infine di sostanze come l'acido lattico. Nelle ultime tappe del processo l'ADP è trasformato in ATP. L'energia contenuta nell'ATP può essere liberata dalla ATP-asi, un enzima legato alla membrana, che utilizza l'energia per espellere protoni (ioni idrogeno) dalla cellula. Come conseguenza il gradiente di protoni aumenta e la cellula diventa alcalina rispetto al suo ambiente. Nelle reazioni inverse, anch'esse catalizzate dall'ATP-asi, i protoni rientrano nella cellula e la energia riguadagnata è utilizzata per formare ATP. Il meccanismo indicato in basso espelle protoni indipendentemente.

venne richiesto di porre il braccio nello spettrometro NMR e si misurarono gli spettri del fosforo 31 ottenuti dal braccio mentre eseguiva un esercizio. I ricercatori avevano già visto che nei muscoli di una persona normale il *p*H diminuisce durante lo sforzo. In un caso il *p*H diminuì da 7,0 a 6,4, quindi ritornò intorno a 7,0 quando l'esercizio fu interrotto. Invece il *p*H del muscolo del paziente partiva da 7,1 e aumentava fino a 7,2. Si osservò anche che negli spettri il picco dovuto alla fosfocreatina diminuiva marcatamente mentre il paziente eseguiva l'esercizio.

Questi risultati sono spiegati dal fatto che, in assenza dell'azione della fosforilasi, le cellule muscolari del paziente hanno poco glucosio disponibile per essere metabolizzato tramite la glicolisi. La glicolisi avrebbe condotto alla formazione di acido lattico, il cui accumulo temporaneo avrebbe diminuito il *p*H delle cellule muscolari normali. La glicolisi avrebbe anche aumentato le riserve di ATP, che sarebbero state disponibili per fornire l'energia necessaria all'esercizio muscolare. Nelle cellule muscolari del paziente l'ATP richiesto era derivato dalla fosfocreatina. La conferma che la fosforilasi del paziente era inattiva fu data più tardi tramite un esame chimico di un campione bioptico del muscolo. Ross e i suoi colleghi osservano che le analisi innocue e che non danneggiano i tessuti, rese possibili dalla NMR, potrebbero essere utili non solo nella diagnosi di malattie come la sindrome di McArdle, ma anche nella valutazione dell'efficacia del trattamento terapeutico.

I dettagli delle reazioni glicolitiche tramite le quali il glucosio è metabolizzato possono essere esaminati nelle cellule viventi per mezzo dell'NMR. Per esempio, tramite la glicolisi ciascuna molecola di glucosio, che ha sei atomi di carbonio, è convertita in due molecole di acido lattico, che ha tre atomi di carbonio. Come ho osservato, la corrispondente acidificazione appare negli spettri NMR del fosforo 31 ottenuti da tessuto muscolare sano come uno spostamento del picco di fosfato inorganico in accordo con la diminuzione del *p*H.

La trasformazione chimica del glucosio può essere seguita più direttamente, tut-

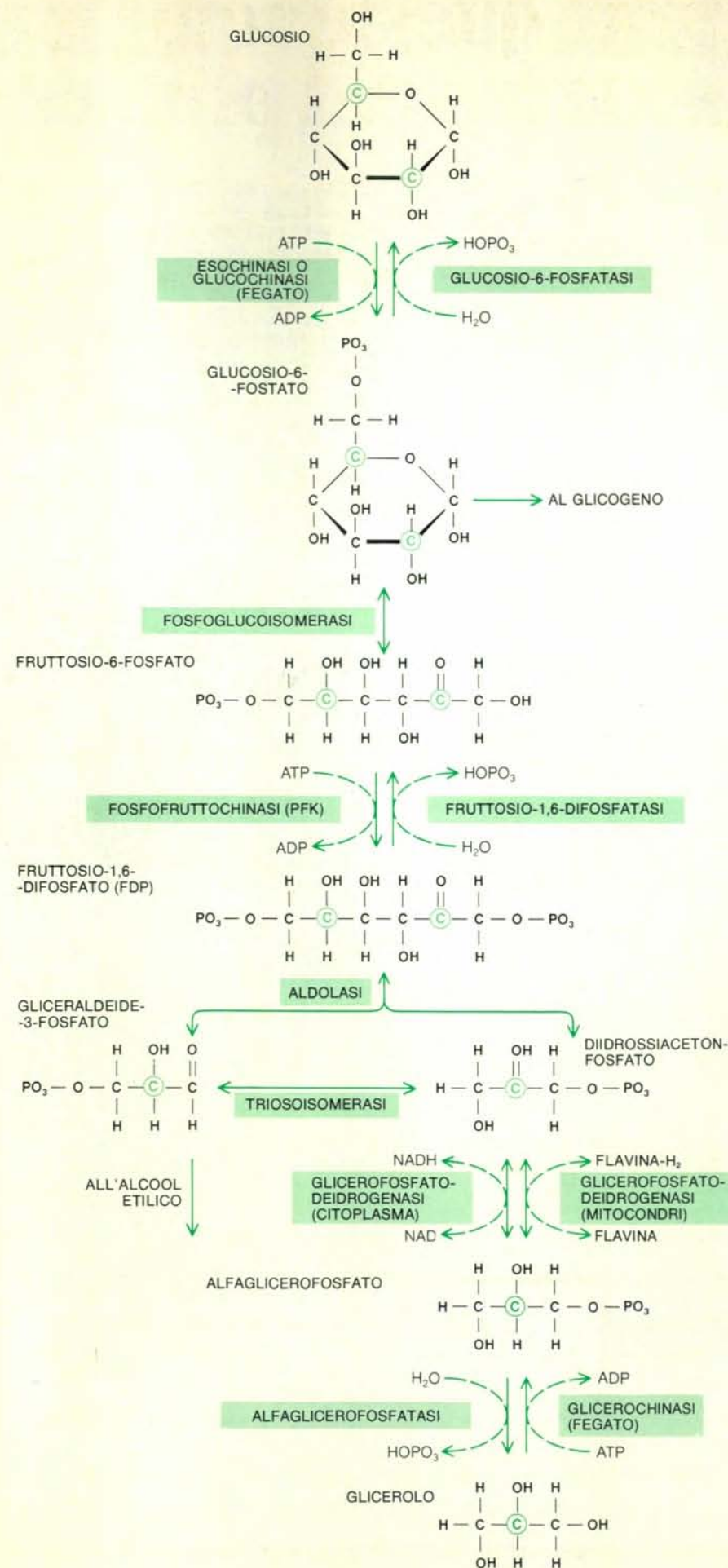
La via glicolitica è la sequenza di reazioni chimiche tramite la quale le cellule metabolizzano glucosio. La sequenza può essere seguita negli spettri NMR aggiungendo glucosio marcato con carbonio 13 alle cellule in esame. Qui il glucosio è stato marcato nella posizione del carbonio C_1 *(cerchietto in colore)*. Poi nel corso della glicolisi il tracciante si sposta nella posizione C_1 del FDP e quindi nel diidrossiacetonfosfato, un composto a tre atomi di carbonio. Se le reazioni avvengono in direzione inversa, una parte sostanziale del tracciante si sposterà nella gliceraldeide-3-fosfato *(cerchietti in nero e in colore)* e quindi nella posizione C_6 del FDP. Infine arriverà alla posizione C_6 del glucosio. Il «rimescolamento» del tracciante nel FDP, misurato tramite l'intensità dei picchi del carbonio 13 negli spettri NMR, costituisce quindi una misura della reazione inversa rispetto alla glicolisi nelle cellule viventi.

L'effetto dell'ormone tiroideo su cellule di fegato di ratto è di aumentare il tasso a cui esse sintetizzano glucosio tramite un processo che è in pratica l'inverso della glicolisi. Uno spettro NMR *(a)* mostra i picchi dovuti al carbonio 13 in sostanze sintetizzate dalle cellule di fegato di ratto dopo che al ratto è stata somministrata alanina marcata con carbonio 13 in posizione C_3. Per ottenere l'altro spettro *(b)* un ratto è stato trattato con ormone tiroideo e poi con alanina marcata. I picchi del carbonio 13 nelle varie posizioni delle due forme di glucosio *(α e β)* sono notevolmente più alti. Anche altri picchi mostrano delle differenze. I picchi segnati con *1, 2, 3* e *4* non sono stati ancora identificati.

tavia, tramite la spettroscopia NMR di cellule alle quali è stato somministrato del glucosio in cui la maggior parte delle molecole sono marcate con carbonio 13 nella posizione del carbonio indicata con C_1. Questa strategia fu portata avanti da J. A. den Hollander nel mio laboratorio presso la Yale University, seguendo i precedenti lavori eseguiti in collaborazione con Ugurbil e T. R. Brown. Lo scopo era quello di misurare i tassi a cui le sostanze intermedie della via glicolitica sono trasformate ciascuna nella successiva entro cellule di lievito, e quindi di correlare questi tassi con le proprietà degli enzimi che catalizzano ciascuna trasformazione.

Da qualche tempo si sa che il tasso a cui le cellule di lievito consumano glucosio è più alto in assenza di ossigeno. L'aumento è detto «effetto Pasteur». Den Hollander, Ugurbil e Brown cominciarono, quindi, a misurare i picchi dovuti al glucosio negli spettri del carbonio 13 ottenuti da una sospensione di cellule di lievito a cui avevano aggiunto le molecole marcate. I risultati mostrarono piuttosto chiaramente che il glucosio era assorbito e consumato due volte più velocemente in assenza di ossigeno.

Il problema seguente era di verificare se l'effetto Pasteur era causato dal cambiamento dei tassi a cui alcuni enzimi agiscono. Una possibilità in particolare sembrò degna di essere esaminata. Dopo alcune tappe della via glicolitica c'è la reazione che trasforma il fruttosio-6-fosfato in fruttosio-1,6-difosfato, o FDP, catalizzata dall'enzima fosfofruttochinasi (PFK). Sappiamo che l'enzima è allosterico: il tasso a cui catalizza la reazione dipende non solo dalla concentrazione del suo substrato (il fruttosio-6-fosfato su cui agisce), ma anche, almeno in vitro, dalla concentrazione di almeno 10 differenti piccole molecole, che per questo sono chiamate effettori. In parte a causa di questa ricchezza di possibili meccanismi di controllo, si era a lungo ritenuto che la PFK fosse regolata nelle cellule. Non c'erano, tuttavia, esperimenti che mostrassero quali effettori regolassero realmente l'enzima nella cellula.

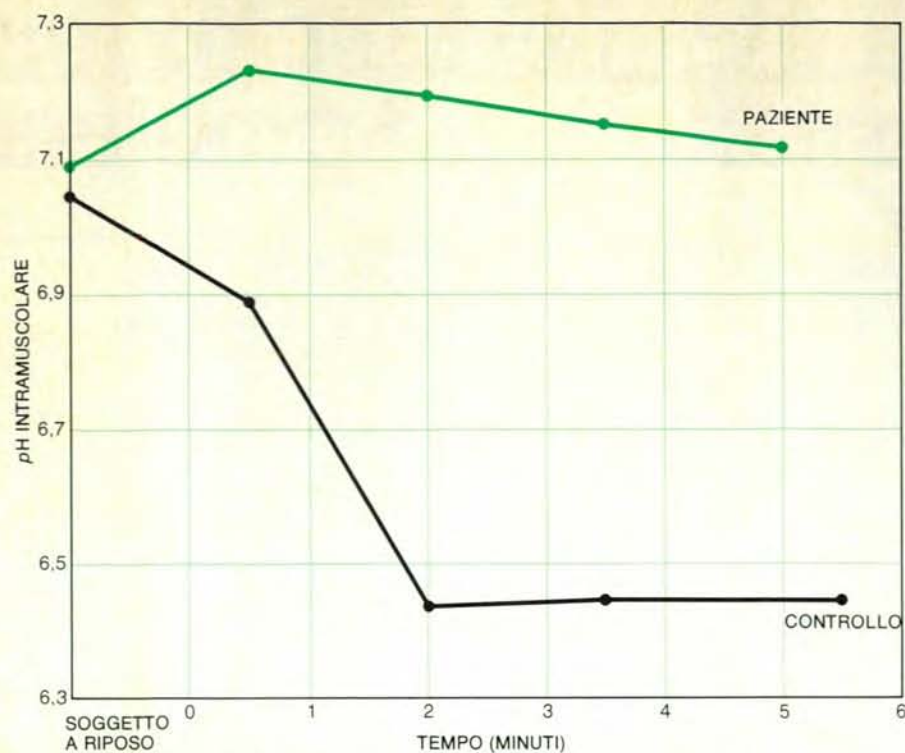
Den Hollander, Ugurbil e Brown utilizzarono gli spettri del fosforo 31 per studiare i precisi cambiamenti causati nelle cellule di lievito dalla presenza di ossigeno. Le loro misurazioni mostrarono un cambiamento significativo del picco del P_i. La sua intensità diminuiva di circa quattro volte e la sua posizione si spostava in modo da mostrare che il *p*H intracellulare era aumentato da 7,04 a 7,45. Allo stesso tempo il picco dovuto al fruttosio-6-fosfato aumentava di parecchie volte e i picchi dovuti all'ATP e all'ADP cambiavano in modo non significativo. I dosaggi in vitro dell'attività della PFK eseguiti recentemente da den Hollander e David Reibstein mostrano che in queste condizioni l'enzima diventa meno attivo.

Eccezionalmente, tuttavia, la concentrazione intracellulare di FDP rimaneva abbastanza costante negli spettri di den Hollander, anche se il tasso a cui il FDP veniva prodotto era marcatamente diminuito. Anche il tasso al quale il FDP veniva consumato doveva essere diminuito, e nella stessa misura. Il flusso del FDP, attraverso questo stadio della glicolisi, può essere dedotto dagli spettri NMR del carbonio 13 tramite un metodo che sfrutta la natura stessa delle reazioni glicolitiche. Nelle prime tappe della glicolisi il tracciante carbonio 13 introdotto da den Hollander nella posizione C_1 del glucosio si ritrova direttamente nella posizione C_1 del FDP. Segue la reazione in cui il FDP

viene scisso in due composti a tre atomi di carbonio che procedono ancora nella via glicolitica. In alternativa, ciascun composto a tre atomi di carbonio può essere trasformato nell'altro composto a tre atomi di carbonio e i due possono poi combinarsi per riformare il FDP. Queste reazioni in direzione inversa «rimescolano» il carbonio 13 così che nel riformare il FDP esso appare nella posizione C_6.

Quindi la competizione tra il flusso catabolico lungo la via glicolitica e il flusso in direzione inversa, in cui il FDP ricompare, determina la distribuzione degli atomi marcati tra le posizioni C_1 e C_6 del FDP. Se c'è uno scarso flusso catabolico lungo la via, ci sarà un ampio intervallo di tempo per riformare il FDP. In questo caso l'intensità del picco dovuto al carbonio 13 in posizione C_6 del FDP sarebbe uguale all'intensità del picco in posizione C_1. Se, d'altro canto, il flusso lungo la via catabolica è rapido, ci sarà poco tempo per rimescolare il tracciante e il picco C_6 sarà assente. Gli spettri di den Hollander mostrano che l'ossigenazione delle cellule di lievito aumenta il rimescolamento. Egli poté in questo modo calcolare che il flusso attraverso il FDP si era ridotto di parecchie volte. Quindi il flusso doveva essere controllato in qualche punto successivo lungo la via metabolica.

L'inversione del flusso lungo la via glicolitica è un oggetto di studio particolarmente interessante in quanto molte cellule invertono il flusso per sintetizzare glucosio quando ne sono carenti. Il FDP è un intermedio in entrambe le direzioni del flusso. L'enzima che forma il FDP lungo la via che si allontana dal glucosio è la PFK; l'enzima che catalizza la reazione inversa, che porta il FDP più vicino di una tappa al glucosio, è la fruttosio-1,6-difosfatasi. È importante sottolineare come la formazione del FDP richieda che un fosfato inorganico sia trasferito dall'ATP al fruttosio-6-fosfato. Nella reazione inversa, invece, il fosfato è semplicemente staccato per idrolisi. Se la reazione avviene in entrambe le direzioni il risultato netto è la perdita di una molecola di ATP in quello che si può considerare un ciclo inutile. Si conoscono molti di questi cicli.

Che cosa controlla un ciclo inutile? Un'ipotesi riguardo al ciclo che coinvolge il FDP era suggerita dal fatto che si sa che il FDP, come molti altri zuccheri, esiste in due forme (alfa e beta) una miscela delle

La marcatura del glicerolo (un prodotto finale del metabolismo del glucosio) rende possibile seguire la via gluconeogenica: la serie di reazioni che ripercorrono in senso inverso la glicolisi. Quando il carbonio situato nel centro del glicerolo è marcato con carbonio 13 *(in colore)*, il tracciante finisce distribuito tra le posizioni C_2 e C_5 del glucosio. Una piccola quantità del tracciante (non rappresentata) finisce nella posizione C_1 del glucosio passando per una serie di reazioni alternative, detta via dei pentosofosfati. Studi NMR della gluconeogenesi mostrano che l'ormone tiroideo aumenta la produzione di glucosio nei ratti ipertiroidei aumentando l'attività dell'enzima glicerofosfatodeidrogenasi.

La diagnosi di una malattia tramite spettroscopia NMR è stata riportata da un gruppo di ricercatori guidati da G. K. Radda e B. D. Ross dell'Università di Oxford. La malattia, sindrome di McArdle, è caratterizzata da astenia dopo esercizio muscolare leggero e deriva dall'inattività dell'enzima fosforilasi, che partecipa alla produzione del glucosio nel tessuto muscolare. I metaboliti del glucosio riducono il *p*H delle cellule; quindi negli spettri NMR la malattia è rivelata da una mancata modificazione del *p*H, come è indicato dalla posizione invariata del picco di fosfato inorganico negli spettri relativi al braccio del paziente mentre eseguiva un esercizio. Ai soggetti esaminati veniva apposto un bracciale per impedire l'arrivo di nuovo glucosio con il sangue arterioso.

quali raggiunge in circa un secondo una distribuzione di equilibrio in cui il 15 per cento delle molecole è in forma alfa e l'85 per cento in forma beta. Studi in vitro mostrano che la fruttosio-1,6-difosfatasi, che avvicina il FDP al glucosio, agisce preferenzialmente sulla forma alfa. L'aldolasi, l'enzima che allontana di una tappa il FDP dal glucosio, agisce preferenzialmente sulla forma beta.

C'era accordo nell'ipotizzare che la cellula vivente controllasse la direzione del flusso attraverso il ciclo inutile controllando la distribuzione delle forme del FDP. La validità di questa ipotesi non potrebbe essere analizzata estraendo il FDP dalle cellule e analizzandone la distribuzione. L'estrazione richiederebbe un periodo di tempo più lungo del secondo necessario per stabilire l'equilibrio. La spettroscopia NMR di *E. coli* invece rende perfettamente realizzabile una analisi di tipo differente. Den Hollander, Ugurbil e Brown misurarono i picchi del carbonio 13 corrispondenti alle forme alfa e beta del FDP nelle cellule, mentre le cellule stavano metabolizzando glucosio. I risultati mostrarono che il FDP era in equilibrio. La direzione del flusso non veniva, quindi, controllata da una condizione di carenza di equilibrio.

Quando il glucosio è metabolizzato nelle cellule di mammifero, i prodotti finali della glicolisi (come l'acido lattico) sono trasportati dal sangue al fegato insieme ai prodotti finali del metabolismo dei grassi (come il glicerolo). Il fegato riceve queste molecole di scarto e le riconverte in glucosio. Il fegato perciò è diventato oggetto di una indagine estensiva sulla via inversa rispetto alla glicolisi: la gluconeogenesi. Paradossalmente i risultati sono troppo abbondanti: lo studio di fegati isolati, di sezioni di fegato e di cellule epatiche rivela una quantità di modi possibili in cui la via metabolica potrebbe essere controllata, ma non fornisce la spiegazione di come essa sia realmente controllata nel fegato degli animali in vivo. Farò un resoconto dei tipi di controllo possibili per fornire una base per discutere i progressi che si stanno facendo tramite la spettroscopia NMR nell'identificare i controlli reali.

Innanzitutto, il tasso di produzione del glucosio nel fegato può essere influenzato dalla concentrazione dei substrati della via gluconeogenica. Un aumentato livello di acido lattico conduce a un tasso maggiore di produzione di glucosio. In secondo luogo, il tasso di produzione di glucosio può essere influenzato dall'apporto di cibo all'organismo. Animali privati di glucosio hanno aumentato i livelli degli enzimi che lo sintetizzano. Terzo, l'attività degli enzimi che sintetizzano glucosio può essere influenzata dal numero di gruppi fosfato legati agli enzimi. Sappiamo che questa fosforilazione, di solito controllata da ormoni a rapida azione come il glucagone, attiva molti enzimi nella cellula. Quarto, parecchi enzimi che sintetizzano glucosio sono controllati da effettori. Quinto, la concentrazione di vari enzimi può essere influenzata da ormoni a lenta azione come quelli sintetizzati dalla tiroide.

Sheila Cohen, che lavorava presso i Bell Laboratories, utilizzò la spettroscopia NMR di composti marcati con carbonio 13 per seguire la via gluconeogenica nel fegato intatto di ratto. Uno dei suoi studi seguì la sintesi del glucosio a partire dal glicerolo, una via costituita da sei reazioni. Varie differenze emersero tra i ratti normali e quelli resi ipertiroidei tramite la somministrazione di alte dosi di ormone tiroideo. Nei ratti ipertiroidei il tasso di produzione del glucosio si mostrò raddoppiato e la concentrazione dell'alfaglicerolfosfato, il composto che segue immediatamente il glicerolo nella via metabolica, apparve diminuita di due volte e mezza.

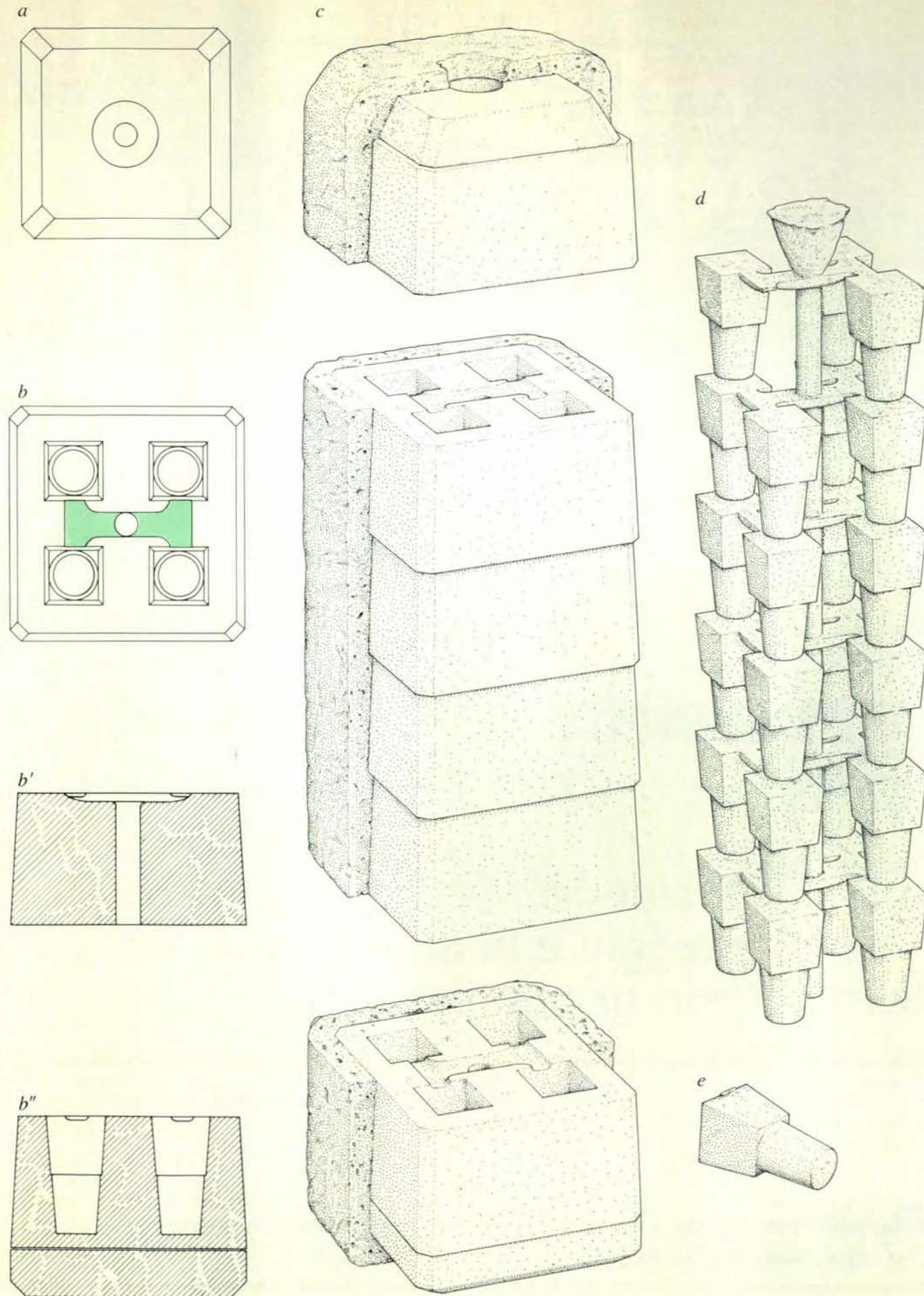
Due enzimi competono per l'alfaglicerolfosfato come substrato. Si tratta di due glicerofosfatodeidrogenasi, ma una si trova nei mitocondri e l'altra nel citoplasma della cellula. Cohen calcolò che sotto l'influenza dell'ormone l'enzima mitocondriale aveva un incremento di quattro volte nel substrato a spese dell'enzima citoplasmatico. La scoperta concorda con i livelli e le attività degli enzimi misurati in estratti di cellule epatiche.

Ci fu anche qualche risultato imprevisto. Per esempio, gli spettri NMR di Cohen mostrarono che l'ormone tiroideo in qualche modo faceva aumentare di parecchie volte il flusso attraverso il ciclo inutile della via gluconeogenica. La base di questi cambiamenti è costituita probabilmente dall'alterazione della concentrazione degli enzimi che è provocata dall'ormone.

Un recente esperimento eseguito da alcuni ricercatori presso gli Oxford Research Systems e nel mio laboratorio a Yale su proposta di Jeffrey R. Alger suggerisce il futuro della spettroscopia NMR. Gli sperimentatori hanno nutrito con glucosio marcato con carbonio 13 un ratto. Quindi negli spettri NMR essi hanno osservato il glucosio scomparire dallo stomaco dell'animale vivo. Per alcune ore la concentrazione di glicogeno nel fegato dell'animale aumentava; il glucosio veniva immagazzinato come glicogeno. Infine il picco di glicogeno scomparve quando il glicogeno marcato fu consumato. Gli spettri NMR hanno fornito ai ricercatori la possibilità di seguire il metabolismo dal principio alla fine nell'organismo dell'animale. Chiaramente la NMR promette di poter osservare con precisione come gli ormoni, la dieta e le malattie influiscono sui processi chimici vitali. Considerando il breve tempo in cui gli spettri NMR sono stati disponibili per studiare il metabolismo, è del tutto ragionevole prevedere che saranno sufficienti pochi anni perché questa tecnica possa rivoluzionare le nostre conoscenze riguardo al procedere delle reazioni biochimiche in vivo negli animali.

La produzione in serie di getti di ghisa nell'antica Cina

Attorno al 500 a.C. i cinesi avevano industrializzato la colata a forme sovrapposte, un'avanzata tecnologia che utilizzando forme multiple sovrapponibili permetteva di ottenere molti getti con una sola colata

di Hua Jue-ming

In Cina la fusione del bronzo risale almeno alla dinastia Xia (XXI-XVI secolo a.C.) ed è una tra le più note tecnologie tradizionali del passato. Non altrettanto è noto, fuori della Cina, lo sviluppo che ebbe in quel paese la produzione in serie di getti di ghisa nel corso del I millennio avanti Cristo. Un gran numero di monete, fibbie per finimenti, parti di carri e di oggetti simili venivano infatti prodotti con la tecnica della colata a forme sovrapposte. Il procedimento consisteva nel sovrapporre l'una sull'altra più forme che, avendo in comune l'imbocco di colata, venivano contemporaneamente alimentate attraverso un unico foro di ingresso del metallo fuso. Era così possibile alimentare fino a una dozzina di forme per volta con il vantaggio di risparmiare metallo e combustibile. Nata tra l'800 e il 500 a.C., la colata a forme sovrapposte è in uso ancora oggi.

Indagini archeologiche recenti hanno portato alla scoperta di resti di fonderie a colata con forme sovrapposte nelle quattro province dello Honan, dello Shantung, dello Shensi e del Chiangsu. Una fornace a ovest della città di Wenxian nello Honan, scoperta or sono quasi dieci anni da Leejing-hua e Tang Wen-xing del Museo provinciale del Honan, è forse il più importante dei ritrovamenti avvenuti. Durante la dinastia degli Han orientali (24-220 d.C.) sorgeva in quel luogo una importante fonderia a colata con forme sovrapposte e la fornace ritrovata ne costituiva la parte principale. Gli scavi condotti in sito hanno permesso di scoprire all'interno della fornace diverse centinaia di serie complete di forme ancora intatte, alcune delle quali capaci di produrre 80 getti per volta. Molte forme erano così ben conservate che fu possibile adoperarle per ottenere getti di ghisa identici a quelli prodotti, per la prima volta, 2000 anni fa.

Prima di descrivere quali nuove conoscenze siano derivate dagli scavi della fornace di Wenxian, sarà utile spendere qualche parola sui fondamenti tecnologici della colata dei metalli e delle fasi che la compongono. La prima fase consiste nella preparazione delle forme, a cui seguono, nell'ordine, la fusione, la colata e la solidificazione del metallo. Nell'intero arco del processo di colata dei getti, la maggior quantità di lavoro viene spesa per la preparazione delle forme. Lungo tutta l'antichità e fino ad arrivare ai tempi moderni, l'argilla è stata la materia prima ideale per costruire le forme. Diffusa e abbondante in tutta la crosta terrestre, è sufficiente poco più di una vagliatura seguita da un lavaggio per ottenere un'argilla fine, esente da impurezze, adatta per la realizzazione di forme.

Mescolando con acqua argilla asciutta, le particelle lamellari molto fini, che la costituiscono, si coprono di un sottile velo liquido che le rende capaci di scorrere l'una sull'altra non appena assoggettate a sforzi di modestissima entità. La tensione superficiale del velo liquido esercita nello stesso tempo un'azione coesiva tra le particelle che ne impedisce la spontanea separazione. Mescolando acqua e argilla in adatte proporzioni si può ottenere di volta in volta un impasto liquido scarsamente denso oppure una pasta appiccicosa. L'argilla asciugando si trasforma in un solido poco resistente. Sottoponendola però all'azione del calore, per esempio cuocendola in un forno, la sua resistenza meccanica aumenta in misura considerevole, grazie all'instaurarsi di un processo di cementazione che salda le particelle d'argilla tra loro. Al pari di altri popoli nel mondo, i cinesi conoscevano bene queste proprietà dell'argilla, materiale impiegato nell'arte ceramica già millenni prima che sorgesse la metallurgia.

Impasti liquidi e paste hanno entrambi un ruolo nella preparazione delle forme. La sezione di una forma può essere ricavata da pasta di argilla foggiata premendovi contro un «modello» fatto di legno o di metallo. Diversamente, un «modello» della forma desiderata può essere spalmato più volte con argilla liquida e messo ad asciugare finché si lascia staccare da quello che è diventato uno spesso guscio di argilla che ne riproduce fedelmente i particolari. In entrambi i casi il «modello» può essere utilizzato per ottenere un numero molto elevato di forme identiche. Bisogna poi lasciare asciugare le sezioni di forma così ottenute, in modo che l'acqua trattenuta dall'argilla possa evaporare, e quindi cuocerle. La cottura non solo aumenta la resistenza meccanica della forma, ma, rimuovendo dall'argilla ogni traccia di acqua residua, elimina il pericolo che la formazione di bolle di vapore provochi difetti superficiali sul

La disposizione di sei forme sovrapposte per getto simultaneo di 24 chiavette cuneiformi di ghisa bianca ad alto tenore di carbonio è indicata nell'illustrazione nella pagina a fronte. In alto, a sinistra, vista in pianta della sommità con l'imbocco di colata *(a)*, entro cui veniva colato il metallo fuso. Sotto, sempre in pianta, cavità *(b)* e i loro collegamenti *(a colori)* con la bocca centrale di colata. Nella vista della sezione verticale *(b')*, in corrispondenza della mezzeria, sono riconoscibili gli attacchi di colata e la continuazione della bocca centrale. In basso *(b'')*, vista della sezione verticale attraverso due delle quattro cavità della forma. Al centro *(c)*, è visibile l'insieme con la forma inferiore e quella superiore esplose. Il rivestimento scuro all'esterno, per metà asportato, è formato da uno strato indurito di fango e paglia, che rendeva solidale l'insieme. A destra *(d)* è visibile il grappolo quale appariva, a solidificazione avvenuta, dopo rottura delle forme. Sotto il grappolo è visibile una chiavetta cuneiforme *(e)* come appariva, separata dall'attacco di colata.

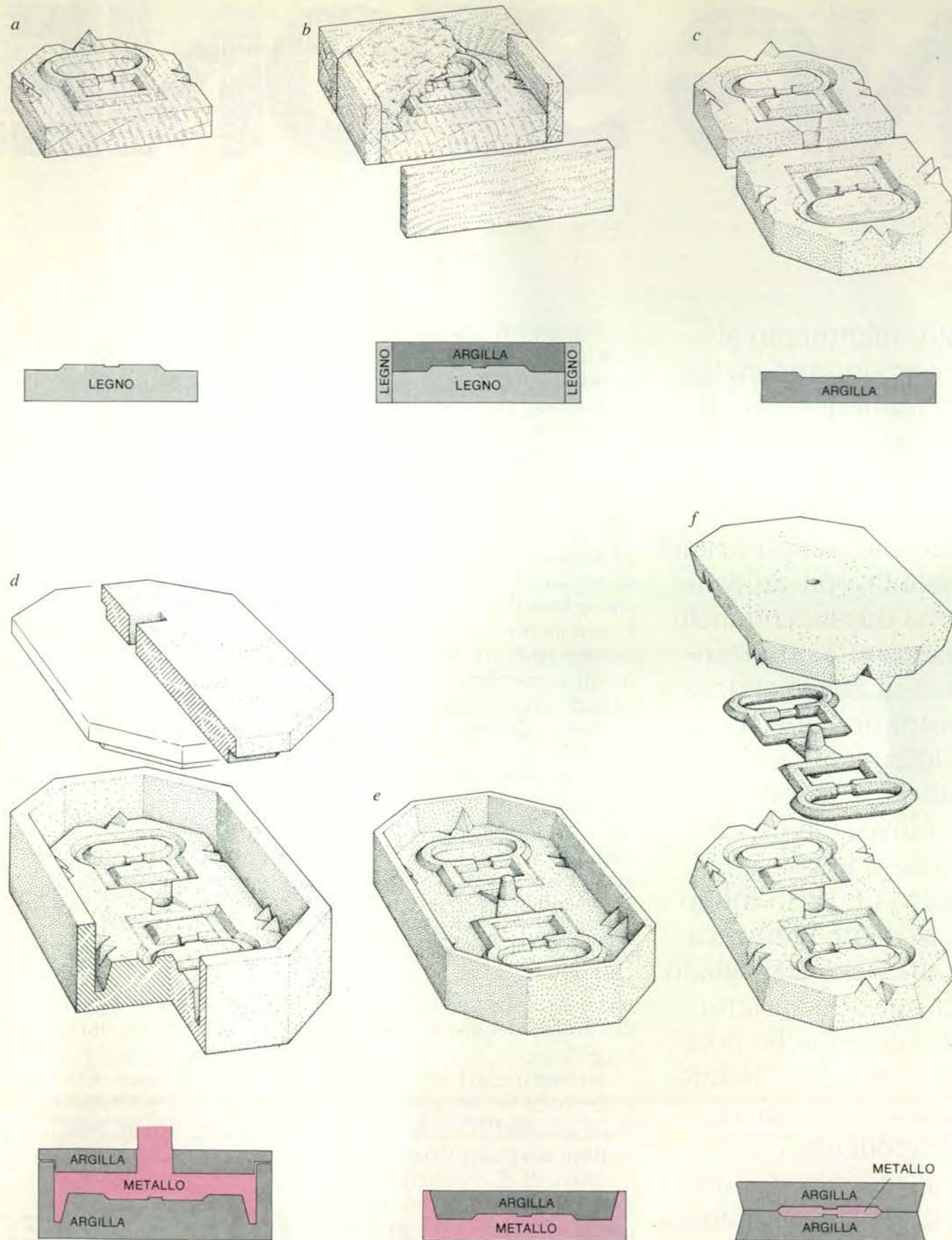
Ritrovamenti di forme sovrapponibili *(pallini bianchi)* si sono avuti in Cina presso Zibo nello Shantung, nei dintorni di Nanchino e nel distretto di Jurong nel Chiangsu. Forme e fornaci *(pallini neri)* sono stati scoperti vicino a Sian nello Shensi e presso Nanyang e Wenxian nel Honan.

metallo colato o deformi, in qualche altro modo, il getto ottenuto. La cottura delle forme può offrire anche un altro vantaggio. Se, come avveniva nella fonderia di Wenxian, le forme sovrapposte venivano tolte dal forno quando la loro temperatura era ancora molto elevata e la ghisa fusa veniva colata in esse prima che avessero il tempo di raffreddare, la solidificazione veniva ritardata e il metallo liquido poteva agevolmente riempire la forma in ogni sua parte, anche la più nascosta. Il ritardo della solidificazione aveva una considerevole importanza per i fonditori di Wenxian perché le loro forme multiple avevano particolarmente stretti i collegamenti di colata, cioè quelle scanalature che permettono alla ghisa liquida proveniente dalla bocca di colata centrale, ovvero dal canale verticale, di raggiungere le cavità di ogni singola forma. Un raffreddamento prematuro della ghisa fusa entro quelle strette scanalature avrebbe comportato scarti di getti incompleti. Non era d'altra parte possibile aggirare il problema semplicemente allargando i collegamenti di colata. La ghisa allora impiegata in Cina non era infatti quella grigia, comune nel mondo occidentale, ma la sua varietà bianca, molto più dura. La durezza di tale ghisa è talmente elevata che il lavoro richiesto per infrangere i collegamenti di colata e sbavare le loro zone di attacco avrebbe considerevolmente ridotto la produttività.

Il primo metallo colato in Cina su vasta scala fu il bronzo, una lega costituita principalmente da rame. I minerali del ferro sono molto più comuni di quelli del rame: non è quindi motivo di sorpresa che il ferro sia divenuto il metallo di impiego corrente in Cina come del resto è avvenuto altrove. In Europa, come pure più a oriente, in India, i primi oggetti di ferro vennero prodotti mediante fusione e successiva fucinatura. Dalla fusione veniva ottenuta una massa spugnosa, ricca di scorie, di ferro quasi puro. La massa, ancora rovente, veniva convertita in metallo utilizzabile a colpi di mazza che ne spremevano la maggior parte delle scorie viscose. Riscaldando il massello così ottenuto in una forgia e fucinandolo su un'incudine si ottenevano oggetti di ferro finiti, quali utensili, armi e ornamenti.

La strada seguita in Cina per ottenere il ferro fu diversa. Gli artigiani cinesi, usando più combustibile e soffiando nei loro forni fusori più aria di quanto non facessero i loro colleghi dell'Europa e dell'Asia occidentale, raggiunsero le condizioni necessarie per portare a completa fusione il minerale ferroso e ottenere un metallo esente da scorie, ricco di carbonio. La presenza del carbonio era importante, perché la temperatura di fusione del ferro diminuisce al crescere del tenore di carbonio in esso contenuto. Il ferro puro non fonde al di sotto dei 1535 gradi centigradi, temperatura che fu possibile raggiungere solo con i forni entrati in uso nel secolo scorso. Però il ferro, quando contiene il 4,3 per cento di carbonio, diviene completamente liquido a 1130 gradi centigradi, temperatura superiore di una ottantina di gradi soltanto a quella necessaria per fondere il bronzo. In Cina, una volta scoperto che la ghisa bianca aveva questa proprietà, non occorse molto a quei fonditori, che già padroneggiavano con maestria la fusione del bronzo, per arrivare a capire che si poteva fare la stessa cosa con il ferro ricco di carbonio, cioè partendo da un minerale molto più abbondante in natura.

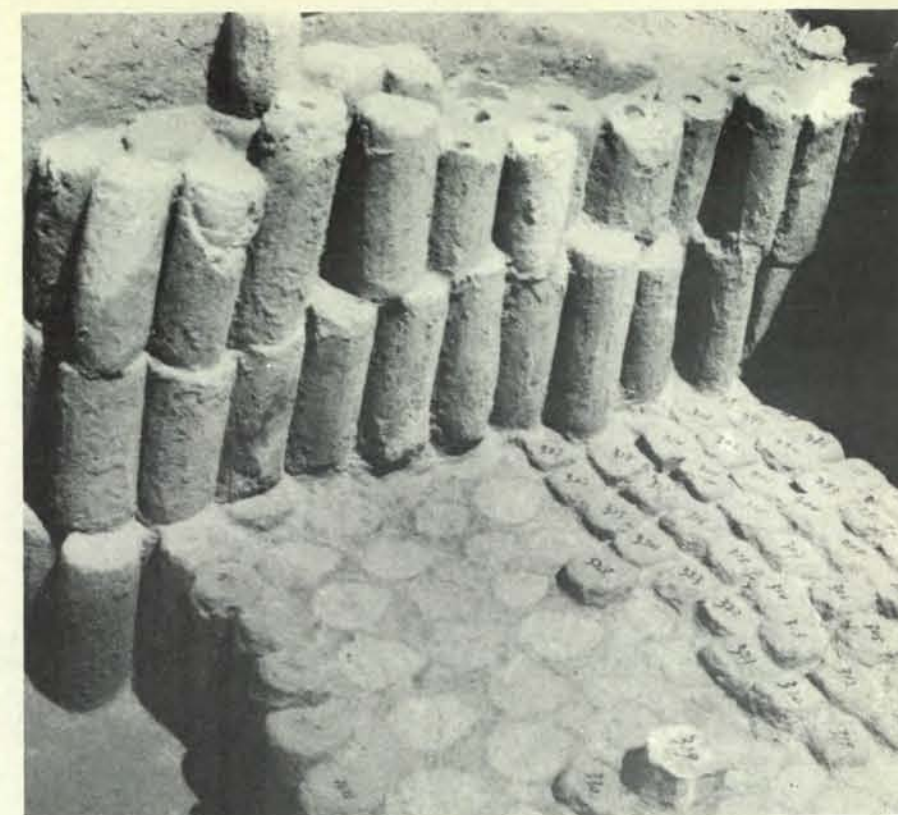
Attorno alla metà del I millennio a.C. la ghisa era ormai comune in Cina. Armi, lame di asce e di scuri, teste di martello, puntali per vomeri e zappe, venivano colati in ghisa e nello stesso modo erano ottenuti articoli ornamentali quali fibbie per cinture e anelli; oggetti di uso domestico quali vasi, pentole e bracieri; equipaggi e morsi per cavalli; parti di assali per carriaggi. Per più di un millennio, dal 300 a.C. al XII secolo si assistette in Cina a un importante sviluppo tecnologico. Si verificò, probabilmente, in quello stesso periodo un aumento della popolazione. Fattore tipico di quella crescita può essere considerato l'aratro. In esso la forza dell'uomo e dell'animale non vengono convogliate soltanto agli scopi delle ordinarie coltivazioni quotidiane, ma anche per dissodare terre vergini. Non va dimenticato, come esempio, che durante l'espansione dell'agricoltura attraverso l'America Settentrionale, nel secolo scorso, erano di ghisa i puntali dei vomeri che solcavano le praterie ancora selvagge. Come altrove, anche in Cina l'uso di vomeri di legno precedette quello di vomeri rinforzati con puntali di ghisa. Non è un azzardo pensare che, grazie alla disponibilità di puntali di ghisa, si poté verificare in Cina verso la fine del I millennio a.C. un mutamento radicale nello sviluppo di terre coltivabili, nella produzione di cibo e, in ultima analisi, nella possibilità di assicurare l'alimentazione che l'espansione demografica richiedeva. Con lo sviluppo dell'agricoltura e dell'artigianato crebbe l'esigenza di trasporti e la richiesta di finimenti e parti di carri divenne forte. Era in questo scenario che la fornace di Wenxian svolgeva il proprio lavoro.

La fornace di Wenxian faceva parte di una fonderia specializzata che produceva soprattutto parti di ghisa per carri trainati da cavalli. Le 500 serie di forme rinvenute a Wenxian erano in grado di fornire getti di 16 tipi diversi, in 36 differenti misure. I getti prodotti comprendevano morsi per cavalli, anelli e fibbie per finimenti, parti usate nel montaggio dei carri

La produzione in serie di forme di argilla era resa più rapida dall'impiego di staffe metalliche. La loro costruzione cominciava scolpendo nel legno la esatta metà dell'oggetto da ottenere, che nel caso illustrato nel disegno è una fibbia *(a)*. Si otteneva poi la riproduzione negativa della metà scolpita stivandole sopra l'argilla *(b)*. In *(c)* sono mostrate due di tali riproduzioni negative con argilla, che successivamente venivano «staffate» insieme *(d)* per costituire la forma definitiva entro cui era ottenuto mediante colata il «modello» di metallo *(e)*. Il modello finito veniva poi impiegato per costruire varie centinaia di forme di argilla che venivano accoppiate per potervi poi fondere le fibbie *(f)*.

come supporti e terminali di assali. Tra le forme per l'ottenimento di getti di altro genere furono rinvenute anche quelle destinate alla produzione di pesi per stadere, un tipo di bilancia che viene utilizzato ancora ai giorni nostri in molti mercati.

Le forme riportate alla luce hanno un grado di uniformità così elevato da far supporre che, sebbene non ne sia stata rinvenuta alcuna, madreforme di metallo venissero impiegate nella loro preparazione. La disponibilità di modelli a copiare, oltre ad assicurare l'uniformità del prodotto, garantiva l'intercambiabilità delle semiforme da accoppiare per costituire una forma completa. Ciò sembra tornasse particolarmente comodo nella produzione di fibbie: alcune delle colonne di forme sovrapposte utilizzate per produrle, formate da 14 strati di 6 impronte ciascuno, erano in grado di fornire 34 getti alla volta. Una larga disponibilità di madreforme offriva un vantaggio ancora maggiore: molti addetti alla formatura potevano operare contemporaneamente e la produttività ne risultava notevolmente accresciuta.

L'analisi dei residui vegetali scoperti nella fornace rivela che le madreforme metalliche venivano cosparse con crusca fine prima che la pasta d'argilla venisse in esse stivata per ottenere le forme con l'impronta del getto. La crusca aveva lo scopo di favorire un nitido distacco delle forme di argilla dalla madreforma di metallo; le forme si lasciavano poi essiccare per circa una settimana in un sotterraneo o in qualche altro punto al riparo dalla luce. Non appena accumulate batterie di forme essiccate all'aria, in quantità sufficiente per una carica della fornace, i fonditori disponevano le forme una sull'altra in successione appropriata, badando che la bocca di colata di ciascuna fosse allineata con la successiva. Per alcuni tipi di forma l'allineamento era ottenuto a incastro, per altri un foro di guida permetteva di infilare dall'alto in basso con una spina tutte le forme. Una volta montate le forme, l'intera colonna veniva rivestita con uno strato di fango misto a paglia triturata (materia prima normalmente usata per i mattoni di fango impastato) e rimesse di nuovo a essiccare. Il rivestimento di fango una volta indurito consentiva di spostare le forme sovrapposte in colonne come un tutto unico. A questo punto le forme erano nuovamente pronte per la cottura.

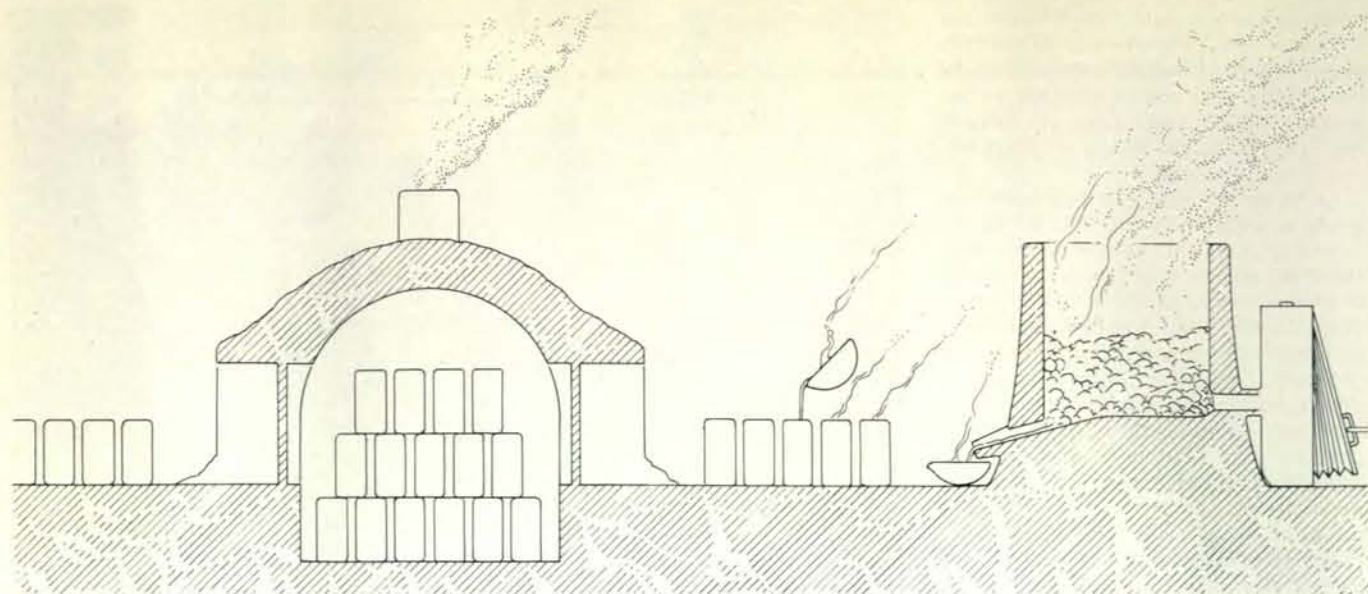
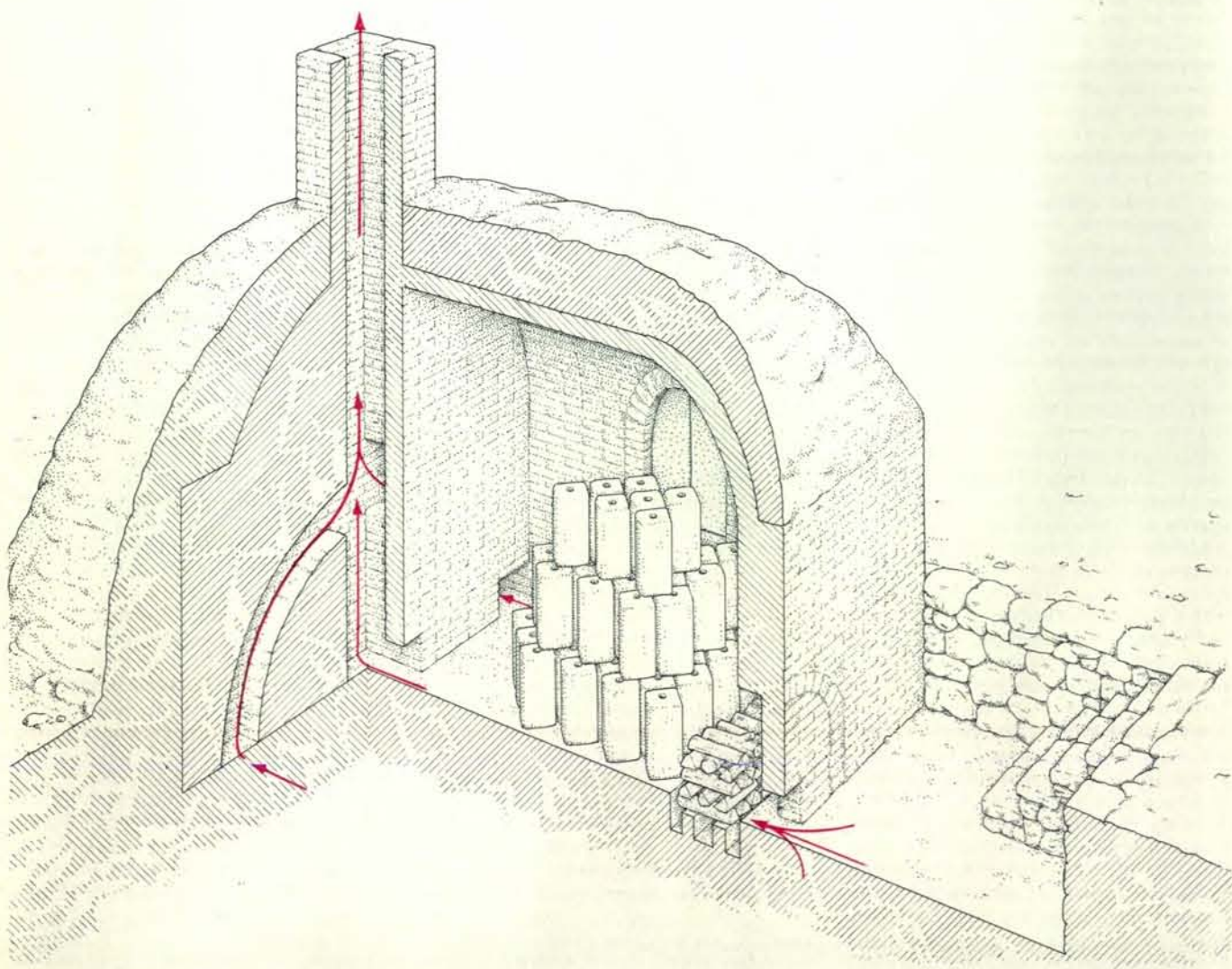
Gli scavi hanno permesso di accertare che il forno di Wenxian era largo 3 e lungo 7,4 metri. Lungo la parete frontale correva una trincea poco profonda in cui veniva acceso il fuoco. Il pavimento della fornace era rivestito di mattoni; tre canne fumarie, nella parte posteriore, fungevano da scarico per i gas caldi. Le batterie di forme sovrapposte, rivestite di fango secco, venivano poi disposte l'una sull'altra e sistemate con cura sul pavimento della fornace in modo che i gas caldi della combustione percorressero lo spazio interposto tra una pila e l'altra prima di raggiungere i condotti di scarico. Si otteneva così

Triplice strato di forme sovrapposte rinvenuto intatto durante lo scavo di Wenxian. Le forme riportate alla luce sono 500; da esse si potevano ottenere 16 getti di tipo diverso in 36 misure.

una uniforme distribuzione di calore nella fornace. A giudicare dai resti di argilla cotta, la temperatura doveva superare i 700 gradi centigradi.

All'epoca della dinastia Han e per molti secoli ancora, il combustibile usato dalla fonderia fu il legno: rami e ceppi per la fornace, carbone di legna per i forni fusori. La continua richiesta di legna da ardere condusse infine a spogliare il paese delle sue foreste (come anche accadde in Europa). In conseguenza di ciò, verso la fine del I millennio d.C. l'estrazione di carbone dalle miniere era diventata in Cina una delle principali attività economiche e da allora fu il carbone a coprire la maggior parte del fabbisogno di combustibile per l'industria.

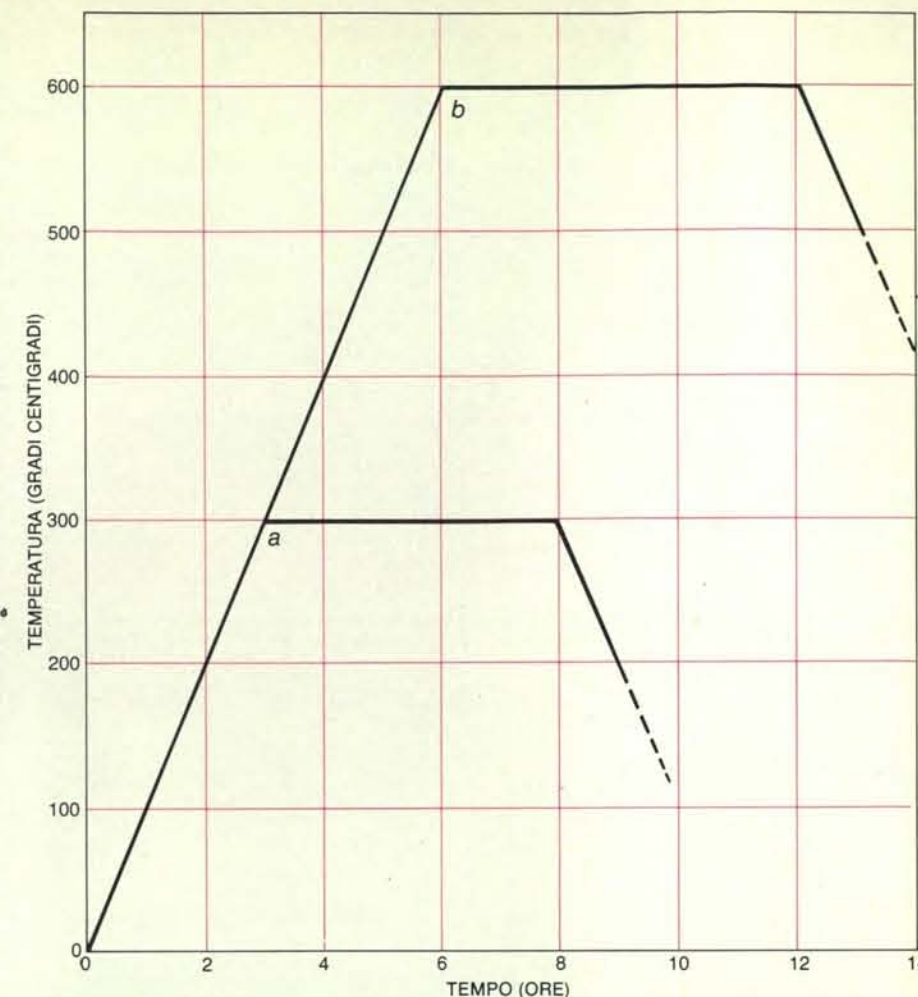
Poiché per mantenere fluido il metallo fuso i fonditori dell'epoca Han colavano la ghisa prima che la loro pila di forme avesse il tempo di raffreddare, è assai probabile che una o più fonderie di ghisa operassero in tandem con la fornace che cuoceva le forme. Solo in tal modo sarebbe stato infatti possibile assicurare la tempestiva disponibilità delle necessarie crogiolate di ghisa liquida, non appena completata la cottura degli stampi. Gli studiosi del sito hanno però dimostrato che, a parte il maggior consumo di combustibile, anche sottoponendo a un nuovo riscaldamento pile, ormai fredde, qualche tempo dopo la primitiva cottura, non si sarebbero incontrate difficoltà, e lo hanno dimostrato riscaldando gli stampi a distanza di quasi 2000 anni! Quell'esperimento ha inoltre permesso di conoscere in qualche misura come probabilmente procedevano i lontani fonditori della dinastia Han.

Queste conoscenze sono state acquisite in seguito alla decisione di sfruttare una parte degli stampi intatti per rimettere in funzione il processo di colata a forme sovrapposte. Alcune serie di forme furono ripulite e poi montate insieme, rivestite di fango e paglia e messe a seccare. Nella prima colata di prova le forme furono mantenute a 300 gradi centigradi per circa cinque ore prima di colare il metallo liquido. A raffreddamento avvenuto la qualità dei getti si rivelò molto scarsa a causa della presenza di grosse porosità e di un ritiro eccessivo. Il riscaldamento delle forme evidentemente non era stato sufficiente. Al secondo tentativo, altre forme vennero portate a 600 gradi centigradi e mantenute a quella temperatura per sei ore piene. A quel punto le forme impilate erano incandescenti. Furono lasciate raffreddare nel forno fino a 300 gradi centigradi e poi estratte; fu quindi colato il metallo (rame o ghisa). Tutti i getti tranne una sola serie risultarono di eccellente qualità. L'eccezione era costituita da un grappolo di chiavette cuneiformi. Tali chiavette avevano forte spessore e le forme non avevano materozze sufficienti. La loro superficie superiore era corrugata, apparentemente a causa di un ritiro nella forma.

La tecnica della produzione di serie a

La fornace per la cottura delle forme sovrapposte di Wenxian *(a sinistra)* avrebbe potuto integrarsi con un forno a cielo aperto per la rifusione dei pani di ghisa *(a destra)* corredato di un mantice per il tiraggio forzato. Nessun forno di questo tipo è stato però finora trovato a Wenxian.

Spaccato della fornace di Wenxian, restaurata dopo lo scavo, con carica di forme sovrapposte pronte per la cottura. Le frecce indicano il percorso dei gas caldi formati dalla combustione della legna *(a destra)* che uscivano dal camino di sinistra dopo aver lambito le pile di forme sovrapposte per garantirne l'uniforme riscaldamento. Nella fornace, larga tre metri e lunga circa sette, la temperatura superava i 700 gradi centigradi.

forme sovrapposte sorse in Cina prima della metà del I millennio a.C. Le forme rinvenute a Wenxian, altamente perfezionate, rappresentano quindi il punto di arrivo di più di cinque secoli di progresso tecnologico nella metallurgia. Poiché per esempio tutte le forme in quella fonderia erano destinate a essere sovrapposte per la colata, esse erano state concepite a superficie di separazione (o montaggio) orizzontale. Ciò poneva un limite naturale al numero di getti più alti che larghi, come ad esempio le chiavette cuneiformi, ottenibili in una sola colata. Nondimeno, controllando rigorosamente l'altezza di ciascuna singola forma, i fonditori di Wenxian riuscivano a colare 24 chiavette alla volta, lavorando con pile di sei forme da quattro getti ciascuna. Nel caso di oggetti quali le fibbie, di altezza inferiore alla rispettiva lunghezza e larghezza, una pila poteva comporsi con un numero di forme molto maggiore e ciascuna forma poteva alloggiare un discreto numero di getti. Uno di tali stampi per fibbie, quantunque non rinvenuto intatto, poteva alloggiare sei fibbie ed era alto solo un decimo della sua lunghezza. Sovrapponendo 20 di tali stampi si sarebbero ottenuti 120 pezzi alla volta.

Fra le forme rinvenute a Wenxian ne è stata trovata una di geometria molto complessa che potrebbe essere stata concepita per la colata di madreforme di metallo da impiegare nella costruzione di forme di argilla per il getto di fibbie appaiate. La preparazione delle madreforme richiedeva molto tempo, ma ciò non comprometteva l'efficienza del ciclo: è stato infatti calcolato che da una sola madreforma si poteva ricavare anche un migliaio di forme prima della comparsa di segni di grave usura.

In altre parti della Cina di quel tempo forme metalliche di semplice fattura venivano impiegate per ottenere direttamente getti anziché forme di argilla. Il loro impiego era limitato alla produzione di utensili di forma elementare come punte di vomere, asce, scuri e cose simili. Serrando l'una contro l'altra due di tali forme elementari di ghisa, veniva ottenuta una conchiglia provvista di foro di alimentazione. Queste forme «permanenti» a parete spessa, robuste, dotate di capacità termica molto superiore a quella dell'argilla presentavano un grande vantaggio: il montaggio della forma completa, la colata e la solidificazione della ghisa liquida, l'estrazione del getto senza attenderne il raffreddamento e il riapprontamento per la successiva colata richiedevano solo pochi minuti. C'è da credere che molto probabilmente sia la produttività sia il costo unitario degli oggetti metallici colati in «conchiglie» fossero più vantaggiosi se confrontati con la fusione a forme di argilla sovrapposte. La colata in conchiglia, tuttavia, venne utilizzata solo per la produzione di oggetti di forma relativamente semplice. Sebbene a Wenxian non sia stata rinvenuta alcuna forma di questo ultimo tipo, l'impiego di forme sia metalliche sia d'argilla nella Cina di due millenni orsono sta a dimostrare che nell'intento di accrescere la produzione delle fonderie dell'epoca Han le buone idee messe in pratica erano più d'una.

Un tentativo di riprodurre la tecnica Han fu compiuto una prima volta *(a)* con cattivo esito riscaldando a 300 gradi centigradi forme sovrapposte e mantenendole a temperatura per poco meno di sei ore. In un secondo tentativo *(b)* la temperatura venne innalzata a 600 gradi centigradi e il tempo di permanenza a sei ore piene, prima di colare a 300 gradi centigradi. I getti così ottenuti risultarono, con una sola eccezione, di qualità elevata.

Come si spiega la fioritura in Cina di una così precoce, sofisticata e ingegnosa metallurgia? I ritrovamenti archeologici indicano concordemente che intelligenza e intraprendenza non facevano difetto all'artigiano del passato. In circostanze favorevoli qualsiasi fonditore, senza distinzione di tempo e di luogo, avrebbe potuto inventare la colata a forme sovrapposte. Cosa rendeva, allora, la Cina diversa? Innanzitutto le tecnologie fondamentali erano già presenti: in Cina, quando ebbe inizio la metallurgia della ghisa, sia la fabbricazione della ceramica sia la fusione del bronzo erano arti note da tempo. In secondo luogo il contesto sociale stimolava fortemente l'innovazione; i vantaggi economico-tecnologici della produzione di serie di utensili, accessori per carri, ornamenti di ghisa a buon mercato in grande quantità, dovevano essere ovvii nel I millennio a.C., diversamente non ne sarebbe stata giustificata la domanda.

Alla luce di tali circostanze la novità delle colate a forme sovrapposte, sebbene di per sé interessante, potrebbe forse suggerire più significativamente la convenienza di esaminare quali circostanze sociali abbiamo potuto propiziare quella innovazione.

Elemento importante della pratica di fonderia durante il primitivo sviluppo della metallurgia cinese, la tecnica a forme sovrapposte non ha cessato di esistere, ma è continuata fino ai nostri giorni. La moderna fonderia di Foshan nella provincia del Kuangtung utilizza tale tecnica per produrre ingranaggi e altre parti di meccanica fine rimarchevoli per precisione e grado di finitura. Le antiche tecnologie scoperte a Wenxian e in altre località non costituiscono per i cinesi soltanto motivo di orgoglioso retaggio, ma li spronano a ulteriori progressi. Trovano così conferma le parole scritte da uno storico della tecnologia, A. B. Wilder: «Studiamo il passato per orientarci nel presente e sperare nel domani. Guardando indietro scrutiamo nel futuro».

TEMI METAMAGICI

di *Douglas R. Hofstadter*

Enunciati simili a virus e strutture autoreplicantesi

Due anni fa, quando scrissi per la prima volta in questa rubrica a proposito degli enunciati autoreferenziali, fui sommerso da una valanga di lettere di lettori interessati al fenomeno dell'autoriferimento nelle sue molteplici forme. Un anno fa ebbi l'occasione di pubblicare alcune di quelle risposte e quell'articolo provocò una seconda ondata di risposte, molte delle quali forniscono nuove illuminanti prospettive sull'autoriferimento. Vorrei qui illustrare le idee di numerosi lettori, iniziando con due che risposero al mio primo articolo con lettere sorprendentemente simili: Stephen Walton di New York e Donald R. Going di Oxon Hill, Maryland.

Walton e Going vedono negli enunciati autoreferenziali una somiglianza con i virus: piccoli oggetti che catturano oggetti «ospiti» più grossi e più autosufficienti, trattenendo gli ospiti con tutti i mezzi per compiere una complicata sequenza di operazioni di replica che portano alla comparsa di nuove copie del virus, le quali sono poi libere di andarsene a catturare altri ospiti. «Enunciati virali», come dice Walton, sono «quelli che cercano di ottenere la loro riproduzione sfruttando le occasioni che sono offerte da entità più complesse».

Sia Walton che Going sono colpiti dalla pericolosità di questi enunciati: l'egoismo con cui invadono uno spazio di idee e si danno da fare, semplicemente producendo copie di se stessi, per prendere possesso di una larga fetta di quello spazio. Perché non procurano di impadronirsi di *tutto* lo spazio? Buona domanda. La risposta dovrebbe essere ovvia per gli studiosi dell'evoluzione: gli enunciati non lo fanno a causa della competizione di altri autoreplicanti. Un tipo di replicante si impadronisce di una regione dello spazio e impara a difenderla dai rivali; si ricava così una «nicchia» nello spazio delle idee.

Questa idea di una lotta evolutiva per la sopravvivenza da parte di idee autoreplicantesi non è nata con Walton e Going, anche se entrambi hanno originali contributi da dare. A quanto ne so, il primo riferimento a questa nozione è in un passo dell'articolo che il neurofisiologo R. W. Sperry scrisse nel 1965 col titolo *Mind, Brain and Humanist Values.* Egli scrive: «Idee causano idee e aiutano nuove idee a evolversi. Esse interagiscono tra loro e con altre forze mentali nello stesso cervello, in cervelli vicini e, grazie alla comunicazione globale, in cervelli molto distanti, estranei. Ed esse interagiscono anche con l'ambiente esterno così da produrre complessivamente un dirompente avanzamento nell'evoluzione, quale mai si è prodotto sulla scena evolutiva, inclusa la comparsa della cellula vivente».

Poco tempo dopo, nel 1970, uscì il libro *Le hazard e la nécessité (Il caso e la necessità,* Mondadori, 1974) del biologo molecolare Jacques Monod. Nell'ultimo capitolo egli scrisse quanto segue a proposito della selezione delle idee:

«Un biologo sente la tentazione di tracciare un parallelo tra l'evoluzione delle idee e quella della biosfera. Benché il regno astratto sia più lontano dalla biosfera di quanto quest'ultima sia lontana dall'universo inanimato, le idee hanno conservato alcune delle proprietà degli organismi. Come questi, esse tendono a perpetuare la loro struttura e a riprodursi; possono anche fondere, ricombinare, scindere il loro contenuto; possono effettivamente evolversi anch'esse e in questa evoluzione la selezione evolutiva deve sicuramente svolgere un ruolo importante. Non azzarderò una teoria della selezione delle idee, ma si può almeno cercare di definire alcuni dei principali fattori in essa implicati. Questa selezione deve necessariamente operare a due livelli: quello della mente stessa e quello dell'esecuzione.

«Il valore esecutivo di un'idea dipende dal cambiamento che induce nel comportamento della persona o gruppo che l'adotta. Il gruppo a cui una certa idea conferisce maggior coesione, maggiore ambizione, maggior fiducia in se stesso, riceve da essa un potenziamento della capacità di espandersi che assicurerà la diffusione dell'idea stessa. La capacità di "aver successo", la misura in cui può essere "messa in pratica", ha poco a che fare con la quantità di verità oggettiva che l'idea può contenere. Un'ideologia religiosa costituisce una salda armatura per una società non in virtù del contenuto, ma perché la sua struttura ne facilita la diffusione. Non si può quindi separare questa capacità che un'idea ha di diffondersi dalla sua capacità di realizzarsi.

«Il "potere di diffusione" - la contagiosità, potremmo dire - delle idee è molto più difficile da analizzare. Diciamo che esso dipende da strutture preesistenti nella mente, tra cui le idee già impiantate dalla cultura, ma indubbiamente anche da certe strutture innate che ci è ben difficile identificare. Quel che è certo, comunque, è che le idee che hanno maggiori possibilità di diffondersi sono quelle che *spiegano* l'uomo assegnandogli il suo posto in un destino immanente, entro cui la sua ansia si dissolve.»

Monod chiama «regno astratto» l'universo delle idee, o ciò che io ho prima chiamato spazio delle idee. Dato che lo descrive in stretta analogia con la biosfera, potremmo chiamarlo «ideosfera».

Nel 1976, il biologo evoluzionista Richard Dawkins pubblicò il suo libro *Il gene egoista* (Zanichelli, 1979), il cui ultimo capitolo sviluppa ulteriormente questo tema. Il nome che Dawkins dà all'unità di replica e selezione nell'ideosfera - il corrispondente nell'ideosfera del gene nella biosfera - è «meme». Proprio come una biblioteca è una raccolta organizzata di libri, così una memoria è una raccolta organizzata di memi. E il substrato in cui i memi crescono - l'analogo del «crogiolo primordiale» da cui è scaturita la vita - è l'ambiente in cui si è sviluppata la cultura umana. Scrive Dawkins:

«Esempi di memi sono i motivi musicali, le idee, gli slogan, le mode, i modi di costruire vasi o archi. Proprio come i geni si propagano saltando da corpo a corpo attraverso lo sperma o le uova, così i memi si propagano saltando da cervello a cervello attraverso un processo che, in senso lato, può essere chiamato imitazione. Se uno scienziato legge o sente una buona idea, la passa ai suoi colleghi e studenti; la cita nei suoi articoli e nelle sue conferenze. Se l'idea prende piede, si può dire che si propaghi da sola, diffondendosi da cervello a cervello. Riassumendo lucidamente una prima stesura di questo capitolo, il mio collega N. K. Humphrey ha scritto: "... i memi andrebbero visti come strutture viventi, non in senso metaforico ma strettamente tecnico. Quando impiantate un meme fertile nella mia mente, letteralmente parassitate il mio cervello, trasformandolo in un veicolo per la propagazione del meme, proprio come un virus può parassitare il meccanismo genetico di una cellula ospite. E non si tratta affatto di un modo di dire - il meme per, ad esempio, 'fede in una vita dopo la morte' si realizza davvero fisicamente, milioni di volte, come una struttura nel sistema nervoso di individui umani in tutto il mondo".

«Consideriamo l'idea di Dio. Noi non sappiamo come sia sorta nel calderone dei memi. Probabilmente si originò molte volte per "mutazione" indipendente. In ogni caso, è certamente vecchissima. Come si replica? Attraverso la parola detta e scritta, con il supporto della grande musica e della grande arte. Perché ha un così alto valore di sopravvivenza? Si ricordi che "valore di sopravvivenza" non significa qui valore per un gene in un insieme di geni ma valore per un meme in un insieme di memi. In realtà la domanda significa: cosa dà all'idea di una divinità la

sua stabilità e capacità di penetrazione nell'ambiente culturale? Il valore di sopravvivenza del meme divinità nell'insieme di memi discende dalla sua grande attrattiva psicologica. Esso fornisce una risposta che può sembrare plausibile a profonde e conturbanti domande sull'esistenza. Suggerisce che le ingiustizie di questo mondo possono venir rettificate in quello successivo. Le "braccia dell'eternità" offrono contro la nostra inadeguatezza un cuscino che, come un placebo, ha efficacia pur essendo immaginario. Queste sono alcune delle ragioni per cui l'idea di Dio è copiata così prontamente dalle successive generazioni di cervelli individuali. Dio esiste, almeno nella forma di meme con alto valore di sopravvivenza, o potere di contagio, nell'ambiente fornito dalla cultura umana.»

Dawkins si preoccupa di sottolineare che non è necessario vi sia una copia esatta di ogni meme, scritta in qualche codice memetico universale, nel cervello di ogni persona. Diverse mutazioni di un meme dovranno contendersi l'una con l'altra, e con altri memi, l'attenzione, cioè le risorse del cervello in termini di spazio e tempo dedicati a quel meme. I memi non devono competere solo per le risorse interne ma anche, potendosi trasmettere attraverso la visione e l'ascolto, per il tempo radiofonico e televisivo, per lo spazio pubblicitario, per il numero di righe negli articoli su giornali e riviste e per lo spazio negli scaffali delle librerie. Inoltre, alcuni memi tenderanno a screditarne degli altri e alcuni gruppi di memi tenderanno ad autorinforzarsi internamente. Scrive Dawkins:

«Denti, artigli, organi interni e organi sensoriali mutuamente adeguati si sono evoluti nei corredi genetici dei carnivori, mentre un insieme stabile di caratteri differente è emerso dai corredi genetici degli erbivori. Avviene qualcosa di analogo nei corredi di memi? Il meme divinità, ad esempio, si associa a qualche altro particolare meme, e questa associazione sostiene la sopravvivenza di ciascuno dei contraenti? Forse potremmo considerare una chiesa organizzata, con la sua architettura, rituali, leggi, musica, arte e tradizione scritta, come un insieme stabile di memi reciprocamente adattati che si sostengono l'un l'altro.

«Per fare un esempio particolare, un aspetto dottrinale che ha avuto grande efficacia nel rinforzare l'osservanza religiosa è la paura del fuoco infernale. Molti bambini e anche qualche adulto credono che soffriranno spaventosi tormenti dopo la morte se non obbediscono alle regole dei preti. Si tratta di una tecnica di persuasione particolarmente ripugnante, che ha causato grandi angosce psicologiche per tutto il Medio Evo e anche ai giorni nostri. Eppure è estremamente efficace. La si direbbe quasi pianificata deliberatamente da una macchiavellica classe sacerdotale esperta in profonde tecniche psicologiche di indottrinamento. Dubito però che i preti fossero così astuti; molto più probabilmente, memi inconsci hanno assicurato la propria sopravvivenza in virtù di quelle stesse caratteristiche di pseudo spietatezza che mostrano i geni di successo. L'idea del fuoco infernale è, molto semplicemente, *autoperpetuantesi* a causa del proprio profondo impatto psicologico. Ha iniziato a collegarsi col meme divinità perché i due si rinforzano l'un l'altro e sostengono la reciproca sopravvivenza nel corredo dei memi.

«Un altro membro del complesso di memi religiosi è chiamato fede. Essa significa fiducia cieca, in assenza di prove, perfino a dispetto dell'evidenza... Nulla è più letale, per un certo tipo di memi, della tendenza a cercare le prove... Il meme della fede cieca assicura il proprio perpetuarsi col semplice espediente inconscio dello scoraggiare l'indagine razionale.

«La fede cieca può giustificare qualsiasi cosa. Se un uomo crede in un'altra divinità, o anche solo usa un diverso rituale per adorare la stessa divinità, la fede cieca può decretare che egli debba morire - sulla croce, sul rogo, infilzato sulla spada di un crociato, colpito da un proiettile in una strada di Beirut, o fatto saltare in un bar di Belfast. I memi della fede cieca hanno i propri modi spietati per propagarsi. E quello che vale per la cieca fede religiosa vale anche per la cieca fede patriottica e politica.»

Quando penso ai memi, spesso mi trovo a immaginare un effimero fluttuare di scintille che saltano da un cervello all'altro gridando «Io, io!». È interessante il modo in cui le lettere di Walton e Going hanno rinforzato questa immagine. Per esempio, Walton inizia con i più semplici enunciati virali che si possano immaginare - «Dimmi» e «Copiami» - e passa subito a più complesse varianti condite di blandizie («Se mi copi, ti prometto di esaudire tre desideri») o di minacce («Dimmi o getterò una maledizione su di te»), nessuna delle quali, egli osserva, è in grado di mantenere la parola. Naturalmente, come rileva, questo può non avere importanza, visto che il solo test finale di vitalità è il riuscire a sopravvivere nel corredo genetico. Tutto è lecito in amore e in guerra - e nella guerra è anche compresa l'eterna battaglia per la sopravvivenza, per i memi nell'ideosfera non meno che per i geni nella biosfera.

Certo ben poche persone sopra i cinque anni cederanno per le ingenue minacce e promesse di questi enunciati. Ma se solo si aggiunge alla frase «nell'aldilà», un numero ben maggiore di persone cadrà nella trappola memetica. Walton osserva che un trucco analogo è usato dalla tipica catena di lettere (o «testo virale»), che «promette salute a quelli che fiduciosamente la riproducono e minaccia rovine a chi non la copia». Ricordate la prima volta che avete ricevuto una lettera di una catena di questo genere? Ricordate la triste storia di «Don Elliott, che ricevette 50 000 dollari ma poi li perse perché aveva rotto la catena»? E la tragica storia del «Generale Welch nelle Filippine, che perse la vita sei giorni dopo aver ricevuto questa lettera perché non la fece circolare, ma prima di morire ricevet-

te 775 000 dollari»? È difficile non farsi un po' suggestionare da questi racconti, anche se alla fine gettate via la lettera con disprezzo.

Io trovo le espressioni di Walton «enunciato virale» e «testo virale» estremamente valide - piccoli memi anch'esse, sicuramente meritevoli, a mio giudizio, di essere replicate qualcosa come 700 000 volte per stampa e chissà quante altre volte oralmente. Lo stesso testo virale di Walton, come potete vedere con i vostri occhi davanti a voi, si è procurato il sostegno di un potente ospite: un'intera organizzazione di stampa, una rivista e un servizio di distribuzione. È saltato a bordo e ora - non appena leggete questo enunciato virale - si propaga follemente per l'ideosfera.

Quest'idea di scegliere un ospite adatto è anch'essa un'importante caratteristica di un'entità virale. Walton la mette in questi termini: «Il recipiente di un testo virale, naturalmente, può dar luogo a grosse differenze. Un virus del mosaico del tabacco che attacca un cristallo di sale è sfortunato e così ci sono alcuni che strappano le lettere di una catena appena le vedono. Un manoscritto inviato a un editore può essere considerato virale, anche se non contiene alcun autoriferimento esplicito, perché sta tentando di assicurarsi la propria riproduzione attraverso un ospite appropriato; lo stesso manoscritto inviato a qualcuno che non ha nulla a che fare con l'attività editoriale può non avere alcuna qualità virale.»

Come conclusione, la lettera di Walton esce graziosamente dalla pagina e confessa direttamente la propria aspirazione: «Infine, io (questo testo) sarei veramente lieto di essere incluso, tutto o in parte, nella sua prossima analisi dell'autoriferimento. Con questa speranza, la prego di scusarmi fin d'ora per averla contagiata.»

Mentre Walton cita Dawkins nella sua lettera, sembrerebbe che Going non lo conosca assolutamente, e questo rende ancor più sorprendente la stretta connessione tra la sua lettera e le idee di Dawkins. Per iniziare, Going ci invita a considerare l'Enunciato *A:* «È nostro dovere convincere gli altri che questo enunciato è vero». Come egli dice: «Se foste abbastanza sciocchi da credere a questo enunciato, cerchereste di convincere i vostri amici che *A* è vero. Se essi fossero altrettanto sciocchi, convincerebbero i loro amici, e così via finché ogni mente umana contenesse una copia di *A*. Quindi *A* è un enunciato autoreplicantesi. Più in particolare, è l'equivalente intellettuale di un virus. Se l'Enunciato *A* dovesse entrare in una mente, prenderebbe il controllo dei suoi meccanismi e li sfrutterebbe per produrre centinaia di copie di se stesso in altre menti.»

«Naturalmente, il problema dell'enunciato *A* è di essere assurdo; nessuno potrebbe crederlo. Ma consideriamo il seguente:

Sistema *S:*
Inizio:
*S*1: Bla.
*S*2: Bla bla.
*S*3: Bla bla bla.
: :
: :
: :
*S*99: bla bla bla bla
bla bla...
*S*100: È vostro dovere convincere gli altri che il Sistema *S* è vero.
Fine.

*S*1-*S*99 vanno intesi come enunciati che, presi insieme, costituiscono un sistema di credenze con un certo grado di coerenza. Se il Sistema *S* preso nel suo insieme fosse convincente, allora l'intero sistema sarebbe autoreplicantesi. Il Sistema *S* sarebbe particolarmente convincente se *S*100 non venisse formulato esplicitamente ma si ponesse come logica conseguenza delle altre idee del sistema.»

Chiamiamo l'*S*100 di Going il «gancio» del Sistema *S*, perché è con questo gancio che il Sistema *S* spera di sollevarsi a un più alto livello di potere. Si noti che di per sé un gancio che dica effettivamente «È vostro dovere credermi» non è una vitale entità virale; per «prendere il volo», deve trascinare con sé una qualche aggiunta, proprio come un rapace ha bisogno di una coda che lo stabilizzi. Una spinta ascensionale pura perde il controllo e si autodistrugge, ma una spinta ascensionale controllata può sollevarsi insieme al suo controllore. Analogamente, *S*100 e *S*1-*S*99 (presi come un insieme) sono in simbiosi: svolgono ruoli complementari e di mutuo sostegno nella sopravvivenza del meme che insieme costituiscono. Ora Going sviluppa ulteriormente il tema:

«Esiste qualche reale sistema di idee che si comporta come il Sistema *S*? Io ne conosco almeno due. Consideriamo il seguente:

Sistema *X:*
Inizio:
*X*1: Chiunque non creda al Sistema *X* brucerà nell'inferno.
*X*2: È vostro dovere salvare gli altri dalla sofferenza.
Fine.

«Se credeste nel Sistema *X*, cerchereste di salvare gli altri dall'inferno convincendoli che il Sistema *X* è vero. Così il Sistema *X* ha un "gancio" implicito che discende dai suoi due enunciati espliciti, quindi il Sistema *X* è un sistema di idee autoreplicantesi. Senza essere empi, si può suggerire che questo meccanismo abbia svolto un ruolo nella diffusione della Cristianità.

«Si noti che il Sistema *X* non è ancora credibile. Ci vuole un bel po' di abilità e di fortuna per produrre un sistema di idee credibile. Una forma più sinistra di autoreplica è l'enunciato *B:* "La borghesia opprime il proletariato". Se credeste in *B*, vorreste liberare il proletariato dalla borghesia. Scoprireste presto che non avete sufficiente forza per raggiungere lo scopo e avreste bisogno dell'aiuto di migliaia di persone che la pensano come voi. Il primo passo per ottenere il loro aiuto sarebbe convincerli che *B* è vero. Quindi un "gancio" discende dall'Enunciato *B* e *B* è pertanto un sistema di idee autoreplicantesi.

«L'Enunciato *B* è solo un caso particolare del generico enunciato *V:* "Il *cattivo maltratta* la *vittima*". Alla parola *cattivo* deve essere sostituito il nome di qualche gruppo reale (capitalisti, comunisti, imperialisti, ebrei, massoni, aristocratici, uomini, stranieri ecc.). *Vittima* deve essere sostituito dalla vittima corrispondente e *maltratta* deve essere precisato come si vuole. Il risultato sarà un sistema di idee autoreplicantesi per le stesse ragioni valide per *B*. Si noti che ciascuna delle sostituzioni suggerite porta a un sistema di idee storicamente attestato. In effetti, è difficile pensare a una forma di estremismo politico che non sia riconducibile a *V* e quindi al "gancio" *S*100. Si esita a spiegare i reali eventi storici nei termini di un così stupido meccanismo, eppure....»

Going porta le sue idee a questa divertente conclusione: «Supponiamo di parodiare la mia tesi proponendo l'enunciato *E:* "Le idee autoreplicantesi cospirano per assoggettare la nostra mente". Questo enunciato "paranoico" è chiaramente un'idea di tipo *V*. Così la tesi sembra descrivere se stessa. Inoltre, se accettiamo *E*, allora dobbiamo dire che questa idea di tipo *V* implica che dobbiamo distruggere tutte le idee di tipo *V*. È il paradosso di Epimenide.»

È interessante notare che tutte queste persone che si sono occupate di queste idee danno esempi che vanno da cose su piccolissima scala, come i motivi musicali accattivanti (Dawkins cita il tema d'apertura della quinta sinfonia di Beethoven) e le frasi (la stessa parola «meme»), fino alla scala grandissima delle ideologie e delle religioni. Dawkins usa il termine «complesso di memi» per questi più ampi agglomerati di memi; io preferisco la singola parola «schema.»

Una ragione è che essa si adatta bene all'uso suggerito dallo psichiatra e scrittore Allen Wheelis nel suo racconto *The Scheme of Things*. Il personaggio centrale è uno psichiatra e scrittore di nome Oliver Thompson e il libro è cosparso dei suoi tetri e meditabondi saggi, inframezzati da luminosi episodi evocatori. Thompson è ossessionato dalla differenza tra, da una parte, «la grezza natura dell'esistenza, disadorna, imprevedibile», che chiama ripetutamente «il modo in cui le cose sono», e, dall'altra parte, «gli schemi delle cose» inventati dagli esseri umani - i modi per trovare un ordine e un senso nel modo in cui le cose sono. Ecco alcune riflessioni di Thompson su quel tema:

«Voglio scrivere un libro... la storia di un uomo la cui vita diviene una metafora dell'intera esperienza dell'uomo sulla Terra. Esso descriverà la sua ricerca in una successione di schemi delle cose, mostrerà la caduta, una dopo l'altra, di ogni configurazione che egli trova, il suo continua-

re a passare a un'altra, sempre nella speranza che lo schema delle cose che trova e di cui al momento si serve *non* sia affatto uno schema delle cose ma la realtà, il modo in cui le cose sono, quindi un assoluto che durerà sempre, nel quale egli possa essere utile, al quale possa contribuire e attraverso il quale possa dare un significato alla sua vita mortale e conquistare così la vita eterna....

«Lo schema delle cose è un sistema d'ordine. Inizia come la nostra concezione del mondo e infine *diviene* il nostro mondo. Noi viviamo nello spazio definito dalle sue coordinate. È vero in modo autoevidente, è accettato così naturalmente e automaticamente che non ci si rende conto di un atto di accettazione. Lo si succhia col latte materno, è continuamente ripetuto a scuola, proclamato dalla Casa Bianca, insinuato dalla televisione, convalidato a Harvard. Come l'aria che respiriamo, lo schema delle cose scompare, diviene semplicemente realtà, il modo in cui le cose sono. È la menzogna necessaria per vivere. Il mondo quale esiste al di fuori di quello schema diviene vago, irrilevante, in buona parte non percepito, alla fin fine inesistente....

«Nessuno schema delle cose ha mai avuto la stessa estensione del modo in cui le cose sono, essendo nello stesso tempo vero per il modo in cui le cose sono. Tutti gli schemi delle cose comportano limitazione e negazione....

«Uno schema delle cose è un progetto di salvezza. La bontà del suo funzionamento dipenderà dal suo campo d'azione e dalla sua autorità. Se è piccola, anche grandi realizzazioni entro il suo progetto non servono per cacciare la morte. Uno schema delle cose può essere grande come la Cristianità o piccolo come la Alameda County Bowling League. Noi cerchiamo lo schema delle cose più grande possibile, non per il desiderio di raggiungere la verità, ma perché più è comprensivo lo schema più vale la sua promessa di allontanare i nostri timori. Se riusciamo a dare un senso alla nostra vita in uno schema cosmico, vivremo nella certezza dell'immortalità. Le caratteristiche di uno schema delle cose che determinano la sua durata e successo sono il suo campo d'azione, l'opportunità che offre di partecipazione e contributo, e la convinzione con cui è considerato vero in modo autoevidente. Il grandissimo successo della Cristianità per un migliaio di anni dipende dall'universalità della sua portata, dal fatto che include e rende conto di tutto, assegnando a ogni cosa un posto adeguato, offrendo a ogni uomo, sia esso principe o accattone, saggio o stupido, il privilegio di lavorare nella vigna del Signore; e dal fatto che è accettato come vero in tutto il mondo occidentale.

«Quando uno schema si modifica per le intromissioni dell'esistenza esterna, esso perde autorità, ha minore capacità di allontanare i timori; i suoi adepti se ne vanno. Alla fine muore, resta solo nella storia, diviene antiquato o primitivo, e infine diviene un mito. Ciò che conosciamo come leggende erano un tempo riproduzioni della realtà. La Chiesa aveva ragione nel fermare Galileo; attività come la sua introducono nello schema delle cose regnante nuove entità che alla fine distruggeranno quello schema.»

Preso nel senso usato da Wheelis, «schema» rimpiazza adeguatamente il «complesso di memi» di Dawkins. Uno schema impone sul mondo un tipo verticale di ordine percettivo, propagandosi implacabilmente, come il Sistema *S* di Going con il suo «gancio». La descrizione che Wheelis fa dell'inadeguatezza di tutti gli «schemi delle cose» a catturare in modo pieno e fedele «il modo in cui le cose sono» richiama la vulnerabilità di tutti i sistemi formali sufficientemente potenti rispetto all'incompletezza o all'incoerenza - una vulnerabilità che dipende da un altro tipo di «gancio»: il gancio gödeliano, che sorge dalla capacità di autoriferimento di tali sistemi, anche se né Wheelis né Thompson citano l'analogia. Torneremo a Gödel tra breve.

Il lettore di *The Scheme of Things* non può non rimanere colpito dalla somiglianza professionale tra Wheelis e il suo protagonista. È impossibile leggere il libro senza sospettare che le concezioni di Thompson riflettano quelle di Wheelis - eppure chi lo può dire? È una spina nel fianco. Ancorpiù stuzzicante è il titolo del libro di Thompson, che Wheelis cita come per caso verso la fine del racconto: è *The Way Things Are* - un sorprendente contrasto con il titolo del libro in cui è contenuto. Ci si chiede: qual è il significato di questa elegante piega letteraria in cui un livello si ripiega su un altro? Qual è il simbolismo di Wheelis all'interno di Wheelis?

Questa contorsione, per cui una cosa (un enunciato, un libro, un sistema, una persona) sembra riferirsi a se stessa ma solo alludendo a qualcosa *che assomiglia* a se stessa, è detta «autoriferimento indiretto». Si può fare qualcosa del genere fissando la propria immagine nello specchio e dicendo: «Quella persona è sicuramente di bell'aspetto!» Questo esempio è molto semplice perché la connessione tra qualcosa e la sua immagine speculare è così familiare e ovvia che non ci sembra esistere alcuna distanza tra referente diretto e indiretto: li consideriamo completamente uguali. Non ci sembra quindi affatto referenziale in modo indiretto.

D'altra parte, questo dipende dalla facilità con cui i nostri sistemi di percezione convertono un'immagine speculare nel suo rovescio, e da altre caratteristiche dei nostri sistemi cognitivi che ci mettono in grado di vedere attraverso numerosi livelli di traduzione senza essere consapevoli dei livelli - come capita quando si guarda attraverso l'acqua e non si vede l'acqua ma quello che giace sul fondo.

Alcuni autoriferimenti indiretti sono naturalmente più sottili di altri. Consideriamo il caso di Sam e Sue, una coppia che conversa ostentatamente dei loro amici Tina e Tim. Sam e Sue hanno dei problemi nella loro relazione e questi problemi sono del tutto analoghi a quelli di Tina e Tim, solo che i sessi sono rovesciati: Sam sta a Sue come Tina sta a Tim nelle loro rispettive relazioni. Quindi, col progredire della conversazione tra Sam e Sue, anche se in superficie essa riguarda i loro amici Tina e Tim, a un altro livello riguarda *in realtà* loro stessi in quanto riflessi in quelle altre persone. È come se parlando di Tina e Tim, Sam e Sue considerino una favola di Esopo che abbia ovvia rilevanza per la loro situazione. Ci sono cose che procedono simultaneamente su due livelli ed è arduo dire quanto l'uno o l'altro dei partecipanti sia conscio dello scambio di messaggi duali - uno che riguarda i loro amici e uno che riguarda loro stessi.

L'autoriferimento indiretto può essere sfruttato nei modi più seri e inattesi. Consideriamo il caso del presidente Reagan, che in una recente occasione di forte tensione tra Russia e America a proposito dell'Iran fece una digressione per ricordare il comportamento del presidente Truman nel 1945, quando questi minacciò senza mezzi termini i Russi alludendo alla possibilità che gli Stati Uniti usassero armi nucleari per contrastare qualsiasi mossa russa in Iran. Col solo richiamare alla memoria quell'occasione, Reagan invitava a stabilire una relazione tra se stesso e Truman e stava quindi pronunciando una minaccia neanche poi tanto velata, anche se nessuno avrebbe potuto rilevare qualcosa di esplicito. Semplicemente, non c'era la possibilità che un essere sufficientemente conscio non individuasse la connessione. La somiglianza tra le due situazioni era troppo palese.

L'autoriferimento si presenta allora davvero in due varietà - diretta e indiretta - o i due tipi sono solo due punti distanti su un continuo? Direi senza esitare che vale la seconda ipotesi. Si può inoltre sopprimere il prefisso «auto», così che il problema diventa relativo al riferimento in generale. L'essenza è semplicemente che una cosa si riferisce a un'altra ogniqualvolta, per un essere conscio, c'è una correlazione sufficientemente convincente tra i ruoli che le due cose sono percepite svolgere in qualche più ampia struttura o sistema. Attenzione, però: per «essere conscio» io intendo una macchina percettiva affamata d'analogie che sta nel mondo grazie alle sue percezioni; non c'è bisogno che sia umano o nemmeno organico. In realtà porterei ancora oltre l'astrazione del termine «riferimento»: la correlazione che il riferimento stabilisce tra sistemi e ruoli non è necessario che sia realmente *percepita* da un simile essere; è sufficiente che la correlazione esista e sia *percepibile* da un simile essere.

Il film *La donna del tenente francese* (tratto da un racconto di John Fowles dallo stesso titolo) fornisce un elegante esempio di ambigui gradi di riferimento. Nel film si sviluppano contemporaneamente due storie, entrambe caratterizzate da un intreccio complesso: una si svolge nell'Inghilterra vittoriana, l'altra nel presente. Già il fatto che siano presenti due intrecci suggerisce, anche se in modo non

diretto, il richiamo a una correlazione. Ma è ben di più quello che viene suggerito. Ci sono analogie strutturali tra i due intrecci: entrambi hanno caratteristiche triangolari, e ciascuna delle due storie è incentrata su un solo lato del triangolo. Per di più, gli stessi due attori ricoprono il ruolo degli amanti in entrambi gli intrecci: li si vede quindi in contesti che si alternano e con tratti caratteriali che pure si alternano. La ragione di questa «coincidenza» è che la storia contemporanea riguarda la lavorazione di un film sulla storia vittoriana.

Con lo svolgersi parallelo delle due storie, si presentano numerose coincidenze che suggeriscono con maggior forza l'opportunità di stabilire una correlazione. È lasciato però allo spettatore il compito di tracciare questa correlazione; non è mai richiamata esplicitamente. Dopo un po' comunque, diventa semplicemente inevitabile. Il bello di questo gioco è la flessibilità lasciata allo spettatore; c'è spazio per ogni licenza artistica nel vedere connessioni, o nel sospettare o perfino inventare connessioni.

Il riferimento indiretto di tipo artistico è molto meno preciso di quello di tipo formale. Quest'ultimo si ha quando due sistemi formali sono isomorfi, cioè quando hanno strutture interne strettamente analoghe, cosicché esiste una corrispondenza biunivoca tra i ruoli nell'uno e quelli nell'altro. In tal caso l'esistenza di un genuino riferimento diventa per noi tanto chiara quanto quella di una persona che parli della propria immagine speculare: assumiamo che questa sia un puro, immediato autoriferimento senza neppure notare i passi di trasformazione mediati dall'isomorfismo. Di fatto, la connessione può sembrare anche troppo diretta perché possa essere chiamata «riferimento»: si potrebbe semplicemente considerarla un'identità.

L'immediatezza percettiva sta alla base di un famoso enunciato della logica matematica, l'enunciato G di Gödel che è considerato autoreferenziale. Tutti accettano l'idea che G parli di un *numero* g; l'acume di Gödel sta nel vedere che il numero g gioca, nel sistema dei numeri naturali, un ruolo strettamente analogo a

```
*
Write
down ten ‘a’s,
eight ‘c’s, ten ‘d’s,
fifty-two ‘e’s, thirty-eight ‘f’s,
sixteen ‘g’s, thirty ‘h’s, forty-eight ‘i’s,
six ‘l’s, four ‘m’s, thirty-two ‘n’s, forty-four ‘o’s,
four ‘p’s, four ‘q’s, forty-two ‘r’s, eighty-four ‘s’s,
seventy-six ‘t’s, twenty-eight ‘u’s, four ‘v’s, four ‘W’s,
eighteen ‘w’s, fourteen ‘x’s, thirty-two ‘y’s, four ‘:’s,
four ‘*’s, twenty-six ‘-’s, fifty-eight ‘,’s,
sixty ‘‘’s and sixty ‘’’s, in a
palindromic sequence
whose second
half runs
thus:
:suht
snur flah
dnoces esohw
ecneuqes cimordnilap
a ni ,s’’‘ ytxis dna s’‘‘ ytxis
,s’,‘ thgie-ytfif ,s’-‘ xis-ytnewt ,s’*‘ ruof
,s’:‘ ruof ,s’y‘ owt-ytriht ,s’x‘ neetruof ,s’w‘ neethgie
,s’W‘ ruof ,s’v‘ ruof ,s’u‘ thgie-ytnewt ,s’t‘ xis-ytneves
,s’s‘ ruof-ythgie ,s’r‘ owt-ytrof ,s’q‘ ruof ,s’p‘ ruof
,s’o‘ ruof-ytrof ,s’n‘ owt-ytriht ,s’m‘ ruof ,s’l‘ xis
,s’i‘ thgie-ytrof ,s’h‘ ytriht ,s’g‘ neetxis
,s’f‘ thgie-ytriht ,s’e‘ owt-ytfif
,s’d‘ net ,s’c‘ thgie
,s’a‘ net nwod
etirW
*
```

L'enunciato autodocumentantesi di Lee Sallow

quello dell'*enunciato* G nel sistema assiomatico in cui è espresso. Questa referenza obliqua, simile a quella di Wheelis, da G a se stesso attraverso la sua «immagine» g viene generalmente accettata come autoriferimento. (Si noti che abbiamo un'ulteriore corrispondenza: G svolge il ruolo di Wheelis e il suo numero di Gödel g quello dell'alter ego di Wheelis, Thompson.)

Le due rappresentazioni astratte che stabiliscono l'autoriferimento di G, ma lo fanno sembrare indiretto, si possono trasformare in una sola rappresentazione, secondo la formula seguente: «Se A si riferisce a B, e B è esattamente come C, allora A si riferisce a C». Ad esempio, possiamo porre che A e C siano Wheelis e che B rappresenti Thompson. Questo rende un «teorema» l'autoriferimento di Wheelis. Ovviamente, questo teorema non ha una dimostrazione rigorosa, dato che la nostra formula deve essere presa *cum grano salis*: essere «esattamente come» qualcos'altro è un'affermazione molto opinabile.

In un contesto formalizzato in cui «è esattamente come» è virtualmente sinonimo di «svolge un ruolo isomorfo a quello di», la formula ha un significato preciso e quindi può giustificare più rigorosamente un teorema. In particolare, se A e C sono uguagliati a G e B è uguagliato a g, allora si avrà: «Se G si riferisce a g e g svolge un ruolo isomorfo a quello di G, allora G si riferisce a G». Dato che sono vere le premesse, anche la conclusione deve essere vera. Secondo questo schema, allora, G è davvero un enunciato autoreferenziale e non una specie di illusione logica ingannevole quanto un quadro di Escher.

L'autoriferimento indiretto comporta l'idea di autoreplica indiretta, in cui un'entità virale, invece di replicarsi esattamente, dà vita a un'altra entità che svolge lo stesso ruolo ma in qualche altro sistema: magari la sua traduzione in francese, magari un elenco di numeri di produzione di tutte le sue parti insieme a buste già intestate contenenti vaglia indirizzati alla ditta produttrice di tutte le parti e una lista di istruzioni che indicano cosa fare delle parti quando arrivano per posta.

La cosa può riuscire familiare ad alcuni lettori: è infatti un riferimento indiretto al gioco di Von Neumann, il rompicapo che ho proposto in questa rubrica nell'aprile scorso per creare un enunciato che si autodescrive, il cui unico contenuto è a livello di parola o di lettera piuttosto che a livello dell'intera frase. Mi sono accorto, ricevendo possibili soluzioni, che molti lettori non hanno capito che cosa si richiedeva da loro. La sfida implicita nel gioco nasceva da un'obiezione alla complessità del «seme» (la parte tra virgolette) nella versione data da Quine del paradosso di Epimenide:

«è falso quando si aggiunge alla sua citazione» è falso quando si aggiunge alla sua citazione.

Per cogliere la stranezza di quanto sopra, immaginate di voler produrre un robot errante per lo spazio che costruisca una copia di se stesso servendosi dei materiali grezzi che incontra nei suoi viaggi. Potreste costruire il robot simmetrico come un essere umano; potreste anche rendere il robot in grado di produrre una copia speculare di qualsiasi struttura incontri sulla sua strada. Immaginate infine di aver programmato il robot per osservare continuamente il mondo alla maniera in cui un'aquila guarda la terra alla ricerca della preda. L'immagine che il robot ha di ciò che cerca è quella di un oggetto identico alla sua metà sinistra. Non è necessario che il robot sia consapevole del fatto che il suo obiettivo sia identico alla sua metà sinistra; può svolgere allegramente la ricerca di un oggetto che sembra semplicemente essere una struttura arbitraria molto complessa. Quando, dopo aver perlustrato l'universo per 17 fantastilioni di anni, il robot finalmente incontra tale struttura, ovviamente attiva la sua abilità di riproduzione speculare e crea una metà destra. Il passo finale consiste nel collegare le due metà così da ottenere la copia. È facilissimo - ammesso che siate disposti ad aspettare 17 fantastilioni di anni.

L'aspetto caratteristico dell'enunciato di Quine è quello che il suo seme ha una complessità pari alla metà della complessità dell'enunciato stesso, il che equivale a dire che è quasi altrettanto complesso. Riprendendo la nostra immagine del robot, ciò che vorremmo avere in un robot che si autoreplica è la capacità di costruirsi proprio dal nulla; vale a dire, ad esempio, estrarre il minerale di ferro, fonderlo, forgiarlo, costruire la bulloneria e le lastre e, infine, assemblare piccole parti di sottounità sempre più ampie fino a ottenere una replica composta realmente di materiali grezzi. Era questo lo spirito del gioco di Von Neumann: volevo una controparte linguistica al «robot autoreplicantesi del secondo tipo».

In particolare si tratta di un enunciato autocostruentesi che costruisce entrambe le proprie metà: il suo nome (tra virgolette) e la sua regola di costruzione, a partire da materia prima linguistica (parole o lettere). Molti lettori hanno fatto l'errore di presentare come seme una lunga successione di parole (o lettere), messe ciascuna tra virgolette, in un ordine specifico e poi di esplicitare quell'ordine nelle regole di costruzione. Avrebbero potuto mettere tra virgolette anche una lunghissima fila ordinata, come ha fatto Quine. L'idea alla base della mia prova, invece, era quella che la struttura dell'oggetto costruito doveva derivare esclusivamente da qualche principio enunciato nella regola di costruzione, non dalla struttura interna del seme.

Proprio come è poco probabile che un robot autoreplicantesi, che vaghi in un qualche ambiente alieno, trovi tutte le sue parti allineate su di un ripiano in ordine di assemblaggio, mentre è più verosimile che debba contare sul suo «cervello» o programma per riconoscere i materiali grezzi dovunque si trovino in modo da raccoglierli e poi assemblarli, così da comporre una copia di se stesso; allo stesso modo l'enunciato desiderato deve trattare i pezzi del seme indipendentemente dall'ordine in cui sono elencati e ciononostante essere in grado di costruire se stesso, nell'ordine giusto, a partire da essi. Così è meglio se racchiudete l'intero seme tra due virgolette invece di mettere tra virgolette le singole parole - bisogna insomma che l'ordine delle parole del seme (o meglio ancora l'ordine delle lettere) non sia esplicitato.

Il seme della soluzione ideale dovrebbe essere un lungo inventario di parti, analogo alla lista degli ingredienti di una ricetta - magari un elenco di 50 *e*, poi 46 *t* e così via. Ovviamente, le lettere non possono rimanere in tale ordine; esse costituiscono semplicemente i materiali grezzi a partire dai quali si deve costruire il nuovo enunciato. Nessuno ha inviato una soluzione il cui seme fosse a un tale livello primordiale. Alcune persone, tuttavia, hanno inviato soluzioni accettabili, anche se non elegantissime, con semi a livello di parole. La prima soluzione corretta che ho ricevuto è quella di Frank Palmer. Palmer ha fornito diverse soluzioni in cui usava maiuscole e minuscole per distinguere, rispettivamente, tra seme e regole di costruzione. Ecco una soluzione leggermente modificata da me:

after alphabetizing, decapitalize FOR AFTER WORDS STRING FINALLY UPPERCASE FIGPBVKXQJZ NONSENSICAL DECAPITALIZE SUBSTITUTING ALPHABETIZING, finally for nonsensical string substituting uppercase words.

Guardiamo come funziona, passo per passo. Dobbiamo ricordare che le istruzioni da seguire sono in lettere minuscole, e che le maiuscole non vanno lette come istruzioni. Perciò le minuscole non sono gli strumenti con cui operare, ma sono piuttosto come il corpo anestetizzato di un paziente che deve essere operato, pronto a svegliarsi e a rimettersi in movimento quando l'operazione è terminata. Partiamo, dunque. Innanzitutto dobbiamo mettere in ordine alfabetico il seme. (Tratto la virgola come se facesse parte della parola che la precede.) Otteniamo il risultato seguente:

AFTER ALPHABETIZING, DECAPITALIZE FIGPBVKXQJZ FINALLY FOR NONSENSICAL STRING SUBSTITUTING UPPERCASE WORDS

Ora dobbiamo togliere le maiuscole. Otterremo qualche parola minuscola: le parole minuscole anestetizzate di cui parlavo prima:

after alphabetizing, decapitalize figpbvkxqjz finally for nonsensical string substituting uppercase words

Bene, ora l'istruzione finale è di individuare una successione di lettere senza senso (è facile: «figpbvkxqjz») e di mettere al suo posto le parole maiuscole (una copia del seme originale). Si ottiene:

after alphabetizing, decapitalize FOR AFTER WORDS STRING FINALLY UPPERCASE FIGPBVKXQJZ NONSENSICAL DECAPITALIZE SUBSTITUTING ALPHABETIZING, finally for nonsensical string substituting uppercase words

E questa è una copia perfetta del nostro enunciato di partenza! Può svegliarsi adesso dalla sua anestesia e riprodursi a sua volta.

Il passo critico era il primo: mettere in ordine alfabetico. Ciò trasforma il seme, ordinato in modo arbitrario, in un enunciato grammaticalmente corretto e sensato, con una operazione meccanica che presume solo una conoscenza dell'ABC. Ma perché no? È ragionevole supporre la conoscenza riguardante solo l'aspetto sintattico della scrittura, senza riferimento al significato convogliato da essa. Questo è il modo in cui gli enzimi nella cellula vivente trattano il DNA e lo RNA che decompongono e ricompongono agendo in modo puramente chimico, senza riferimento ai «significati» sottostanti. Esattamente come nel funzionamento della cellula le valenze e le affinità chimiche sono considerate come qualcosa di dato, così i fatti alfabetici e tipografici sono considerati come dati nel Gioco di Von Neumann.

Nella soluzione di Palmer il seme si presenta ordinato secondo la lunghezza delle parole, ma ciò è inessenziale. Ogni ordinamento casuale avrebbe raggiunto lo scopo, ed è questo il punto sfuggito a molti lettori. Un'altra elegante soluzione è stata mandata da Martin Weichert di Monaco. Eccola:

Alphabetize and copy in quotes «in and copy quotes Alphabetize»

Funziona come quella di Palmer e consiste di un seme la cui struttura interna è irrilevante per la riproduzione. Weichert ha mandato una soluzione palindroma in esperanto, in cui l'ordinamento flessibile delle parole del linguaggio gioca un ruolo fondamentale. Michael Borowitz e Bob Stein di Durhan, N. C., hanno mandato una soluzione simile a quella di Palmer.

Infine il vincitore della medaglia d'oro dello scorso anno per gli enunciati autodocumentantisi, Lee Sallows, era leggermente piccato per la mia osservazione sull'oro della sua medaglia, leggermente sbiadito dal fatto che non aveva prestato sufficiente attenzione alla distinzione tra uso e menzione. Ciò evidentemente lo ha spinto a elaborare un enunciato autodocumentantesi ancor più elaborato. Anche se non si autoreplica, dato che non precisa esplicitamente la propria costruzione, è un'altra meravigliosa gemma di Sallow e penso che quest'anno ammetterò che l'oro della sua medaglia non è sbiadito.